


```
Db 181 GGGTCAGATTATGTGCAAGAACCTTTGGCGCTTTCTTGCGCAGCAAGTACCATGGGGTT 240
Qy 241 GTGCTCGAAGTGTCTACCGAGAGACAGTCAATGCAAGGAAAGATGAGGGCTTTAGCTGAGCA 300
Db 241 GTGCTCGAAGTGTCTACCGAGAGACAGTCAATGCAAGGAAAGATGAGGGCTTTAGCTGAGCA 300
Qy 301 AGCCACTCAGGCTGCTCAGGCGCAATCTGCAAGCTGTGCTGTTCAAGCCCCCGGCTCC 360
Db 301 AGCCACTCAGGCTGCTCAGGCGCAATCTGCAAGCTGTGCTGTTCAAGCCCCCGGCTCC 360
Qy 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAGAACATGATTGTGCGGCATCAATC 420
Db 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAGAACATGATTGTGCGGCATCAATC 420
Qy 421 TTCCAGCTATCAACAAGACCTGTTACCCCGCTGCAAGTGCCTCTCAGGCGATGAATC 480
Db 421 TTCCAGCTATCAACAAGACCTGTTACCCCGCTGCAAGTGCCTCTCAGGCGATGAATC 480
Qy 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGCAAGGAGGTG 540
Db 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGCAAGGAGGTG 540
Qy 541 TGGATTGACAGGATTTAAAGTGTGCTGTGCGCAACCTCTAATGCGCTTTACATCGGTAATC 600
Db 541 TGGATTGACAGGATTTAAAGTGTGCTGTGCGCAACCTCTAATGCGCTTTACATCGGTAATC 600
Qy 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGCATTGCGAAAGC 660
Db 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGCATTGCGAAAGC 660
Qy 661 TAACTCTCTGTGCGGCGGAGAGGTTGTCAAGTTTGTATGAGCATCCGTTAACTTTT 720
Db 661 TAACTCTCTGTGCGGCGGAGAGGTTGTCAAGTTTGTATGAGCATCCGTTAACTTTT 720
Qy 721 CTGCCGACGATTAAAGCTTCATACATTGAGTAATCTCTAATCTTTCTTTATGAGAG 780
Db 721 CTGCCGACGATTAAAGCTTCATACATTGAGTAATCTCTAATCTTTCTTTATGAGAG 780
Qy 781 AGCGAGTGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTGCATCAAGAGCTCGCC 804

RESULT 2
AX288141 804 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 12 from Patent WO0177311.
ACCESSION AX288141
VERSION AX288141.1 GI:17049843
KEYWORDS
SOURCE .
ORGANISM Physcomitrella patens
            Physcomitrella patens
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE 1
            da Costa Silva,O., Bohnert,H.J., Van Thielén,N. and Chen,R.
            Transcription factor stress-related proteins and methods of use in
            plants
            Patent: WO 0177311-A 12 18-OCT-2001;
            BASF Plant Science GmbH (DE)
FEATURES
            source
            1..804
            /organism="Physcomitrella patens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:3218"

ORIGIN
Query Match 100.0%; Score 804; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 5,6e-236;
Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATCCGGGACACAGTCCCGCTTAAAGTGTGTGTCATTAGTGTGTTGCGAAGCTGAAGC 60
1
1
```

```
Db 1 ATCCGGGACACAGTCCCGCTTAAAGTGTGTGTCATTAGTGTGTTGCG 60
Qy 61 CTGAGCGAGATTTGACAGATTTTCTCATACGCTTCTGATTAGAAAGA 120
Db 61 CTGAGCGAGATTTGACAGATTTTCTCATACGCTTCTGATTAGAAAGA 120
Qy 121 TAGTGTGTTAAAGATGCGCACCGAGCGGTGTCTCAGAGAGACGACTCG 180
Db 121 TAGTGTGTTAAAGATGCGCACCGAGCGGTGTCTCAGAGAGACGACTCG 180
Qy 181 GGGTCAGATTATGTGCAAGAACCTTTGGCGCTTTCTTGCGCAGCCAGCT 240
Db 181 GGGTCAGATTATGTGCAAGAACCTTTGGCGCTTTCTTGCGCAGCCAGCT 240
Qy 241 GTGCTCGAAGTGTCTACCGAGAGACAGTCAATGCAAGGAAAGTGAAGGCT 300
Db 241 GTGCTCGAAGTGTCTACCGAGAGACAGTCAATGCAAGGAAAGTGAAGGCT 300
Qy 301 AGCCACTCAGGCTGCTCAGGCGCAATCTGCAAGCTGTGCTGCTGTTCAAG 360
Db 301 AGCCACTCAGGCTGCTCAGGCGCAATCTGCAAGCTGTGCTGCTGTTCAAG 360
Qy 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAGAACATGATTGTG 420
Db 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAGAACATGATTGTG 420
Qy 421 TTCCAGCTATCAACAAGACCTGTTACCCCGCTGCAAGTGCCTCTCAG 480
Db 421 TTCCAGCTATCAACAAGACCTGTTACCCCGCTGCAAGTGCCTCTCAG 480
Qy 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGCG 540
Db 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGCG 540
Qy 541 TGGATTGACAGGATTTAAAGTGTGCTGTGCGCAACCTCTAATGCGCTTTA 600
Db 541 TGGATTGACAGGATTTAAAGTGTGCTGTGCGCAACCTCTAATGCGCTTTA 600
Qy 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGCG 660
Db 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGCG 660
Qy 661 TAACTCTCTGTGCGGCGGAGAGGTTGTCAAGTTTGTATGAGCATCC 720
Db 661 TAACTCTCTGTGCGGCGGAGAGGTTGTCAAGTTTGTATGAGCATCC 720
Qy 721 CTGCCGACGATTAAAGCTTCATACATTGAGTAATCTCTAATCTTTCTTC 780
Db 721 CTGCCGACGATTAAAGCTTCATACATTGAGTAATCTCTAATCTTTCTTC 780
Qy 781 AGCGAGTGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTGCATCAAGAGCTCGCC 804

RESULT 3
AR452942 1531 bp DNA linear 2004
DEFINITION Sequence 4 from patent US 6677504.
ACCESSION AR452942
VERSION AR452942.1 GI:42685089
KEYWORDS
SOURCE .
ORGANISM Unknown.
            Unclassified.
REFERENCE 1 (bases 1 to 1531)
            da Costa e Silva,O., Bohnert,H.J., van Thielén,N. a
            Transcription factor stress-related proteins and me
            Plants
            Patent: US 6677504-A 4 13-JAN-2004;
            BASF Plant Science GmbH; Ludwigshafen;
            DEX;
            Location/Qualifiers
```


OY 788 CCGATCAGA 797
| | | | |
Db 1308 CCGATCAGA 1317

RESULT 5
AC161406

LOCUS
DEFINITION AC161406 112882 bp DNA linear HTG 13-MAY-2005
Medicago truncatula chromosome 2 clone mth2-83p17, *** SEQUENCING
IN PROGRESS ***, 32 unordered pieces.

AC161406
AC161406.1 GI:63987046
HTG: HTGS PHASE1, HTGS ACTIVEFIN.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 112882)
Town,C.D., Tallon,L.T., Arbogast,T., Althoff,R., Hine,E.,
Moneghan,E., Smith,S.A., Utecherack,T., Feldblyum,T., Koo,H. and
Cheung,F.
Medicago truncatula BAC genomic sequence
2 (bases 1 to 112882)
Town,C.D.
Direct Submission
Submitted (13-MAY-2005) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 10642: contig of 10642 bp in length
10643 10742: gap of unknown length
10743 21453: contig of 10711 bp in length
21454 21553: gap of unknown length
21554 23120: contig of 1567 bp in length
23121 23220: gap of unknown length
23221 28798: contig of 5578 bp in length
28799 28898: gap of unknown length
28899 30125: contig of 1227 bp in length
30126 30225: gap of unknown length
30226 39880: contig of 9655 bp in length
39881 39980: gap of unknown length
39981 41417: contig of 1437 bp in length
41418 41517: gap of unknown length
41518 42809: contig of 1292 bp in length
42810 42909: gap of unknown length
42910 44087: contig of 1178 bp in length
44088 44187: gap of unknown length
44188 47787: contig of 3600 bp in length
47788 47887: gap of unknown length
47889 48891: contig of 1004 bp in length
48892 48991: gap of unknown length
48992 50482: contig of 1491 bp in length
50483 50582: gap of unknown length
50583 51672: contig of 1050 bp in length
51673 51772: gap of unknown length
51773 52840: contig of 1068 bp in length
52841 52940: gap of unknown length
52941 70023: contig of 17083 bp in length
70024 70123: gap of unknown length
70124 71565: contig of 1442 bp in length
71566 71665: gap of unknown length
71666 72803: contig of 118 bp in length
72804 72903: gap of unknown length
72904 74559: contig of 1656 bp in length

* 74560 74559: gap of unknown length
* 74560 79544: contig of 4885 bp in length
* 79545 79644: gap of unknown length
* 79645 90908: contig of 11264 bp in length
* 90909 91008: gap of unknown length
* 91009 92604: contig of 1596 bp in length
* 92605 92704: gap of unknown length
* 92705 94232: contig of 1528 bp in length
* 94233 94332: gap of unknown length
* 94333 95349: contig of 1017 bp in length
* 95350 95449: gap of unknown length
* 95450 100396: contig of 4947 bp in length
* 100397 100496: gap of unknown length
* 100497 101862: contig of 1366 bp in length
* 101863 101962: gap of unknown length
* 101963 103478: contig of 1516 bp in length
* 103479 103578: gap of unknown length
* 103579 105211: contig of 1633 bp in length
* 105212 105311: gap of unknown length
* 105312 106455: contig of 1144 bp in length
* 106456 106555: gap of unknown length
* 106556 107581: contig of 1026 bp in length
* 107582 107681: gap of unknown length
* 107682 110219: contig of 2538 bp in length
* 110220 110319: gap of unknown length
* 110320 111674: contig of 1355 bp in length
* 111675 111774: gap of unknown length
* 111775 112882: contig of 1108 bp in length.

FEATURES
source
1..112882
Location/Qualifiers
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/chromosome="2"
/clone="mth2-83p17"
10643..10742
/estimated_length=unknown
21454..21553
/estimated_length=unknown
23121..23220
/estimated_length=unknown
28799..28898
/estimated_length=unknown
30126..30225
/estimated_length=unknown
39881..39980
/estimated_length=unknown
41418..41517
/estimated_length=unknown
42810..42909
/estimated_length=unknown
44088..44187
/estimated_length=unknown
47788..47887
/estimated_length=unknown
48892..48991
/estimated_length=unknown
50483..50582
/estimated_length=unknown
51673..51772
/estimated_length=unknown
52841..52940
/estimated_length=unknown
70024..70123
/estimated_length=unknown
71566..71665
/estimated_length=unknown
72804..72903
/estimated_length=unknown
74560..74559
/estimated_length=unknown
79545..79644
/estimated_length=unknown

gap 90909. 91008 /estimated_length=unknown
gap 92605. 92704 /estimated_length=unknown
gap 94233. 94332 /estimated_length=unknown
gap 95350. 95449 /estimated_length=unknown
gap 100397. 100496 /estimated_length=unknown
gap 101863. 101962 /estimated_length=unknown
gap 103479. 103578 /estimated_length=unknown
gap 105212. 105311 /estimated_length=unknown
gap 106456. 106555 /estimated_length=unknown
gap 107582. 107681 /estimated_length=unknown
gap 110220. 110319 /estimated_length=unknown
gap 111675. 111774 /estimated_length=unknown
gap /estimated_length=unknown
ORIGIN
Query Match 13.1%; Score 105; DB 14; Length 112882;
Best Local Similarity 67.7%; Pred. No. 2e-20;
Matches 147; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 509 CCCATCGATGCGGATCTTGACGAGAGCGGTGATGACAGGATTTAAGTCCCTGT 568
DB 84612 CCTAAGCGGTGACGACACTGCAATTAAGCGGTGTTTGAAGATTTATGTCATGT 84671
QY 569 GCGAAGCTTACTGCGCTTACTCGGTACTCGGACAAACACTTGCACATATGACTAC 628
DB 84672 GGTATCTTACTGCTGCTGCTACTGCTACTGAGACAGCATGCTCCCATTTGATTAC 84731
QY 629 AAGCCGACGAGGAGAGCGATTTGCGAAAGCTAATCTCTTGTCTGCGCCGAGAAAGTT 688
DB 84732 CCACACTGCTGACGATGATCAATAGCAAAAGCCGTGTGTGAAGGCTGAAAGCTT 84791
QY 689 GTCAGTTTGTATGAGCATCCGTTAAGCTTTCTGCG 725
DB 84792 GACAAAGATCTAAGTTTGTGTTTAAAGTTTCATGTC 84828
RESULT 6
AP007400/c 102653 bp DNA linear HTG 28-DEC-2004
LOCUS AP007400 Lotus corniculatus var. japonicus chromosome 3 clone UjT31F23, ***
DEFINITION SEQUENCING IN PROGRESS ***, 37 unordered pieces.
ACCESSION AP007400
VERSION AP007400.1 GI:56805716
KEYWORDS HTG; HTGS PHASE1.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus
Bukariota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaceae;
Lotus.
1
REFERENCE
AUTHORS Kaneko, T., Asamiizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE Structural Analysis of a Lotus japonicus Genome. XI. Sequence
FEATURES and Mapping of Nine hundred twenty-one TAC Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 102653)
AUTHORS Sato, S.
TITLE Direct Submision
JOURNAL Submitted (26-OCT-2004) Shuei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research, 2-5-7 Kazusa-kamatori, Kisarazu,
Chiba, 292-0818, Japan (E-mail:issato@kazusa.or.jp,
URL: <http://www.kazusa.or.jp/>, Tel:81-438-52-3935(ex.2337),

COMMENT

Fax:81-438-52-3934)
* NOTE: This is a 'working draft' sequence. It con
* consists of 37 contigs. The true order of the p
* is not known and their order in this sequence r
* arbitrary. Gaps between the contigs are represent
* runs of N, but the exact sizes of the gaps are u
* This record will be updated with the finished se
* as soon as it is available and the accession num
* be preserved.
1
772 771: contig of 771 bp in length
872 871: gap of unknown length
1868 1867: contig of 996 bp in length
1968 1967: gap of unknown length
2909 2908: contig of 941 bp in length
3009 3008: gap of unknown length
3973 3972: contig of 964 bp in length
4073 4072: gap of unknown length
5829 5828: contig of 1756 bp in length
5929 5928: gap of unknown length
7171 7170: contig of 1242 bp in length
7271 7270: gap of unknown length
8999 8998: contig of 1728 bp in length
9099 9098: gap of unknown length
10095 10095: contig of 997 bp in length
10966 10965: gap of unknown length
10966 10965: gap of unknown length
11820 11819: contig of 1624 bp in length
11920 11919: gap of unknown length
13526 13525: contig of 1606 bp in length
13526 13525: gap of unknown length
13626 13625: contig of 1470 bp in length
15096 15095: gap of unknown length
15196 15195: gap of unknown length
17491 17491: contig of 2236 bp in length
17492 17491: gap of unknown length
17592 17591: gap of unknown length
19659 19658: contig of 2067 bp in length
19759 19758: gap of unknown length
21398 21397: contig of 1639 bp in length
21497 21497: gap of unknown length
21498 21497: gap of unknown length
23876 23875: contig of 2378 bp in length
23976 23975: gap of unknown length
25652 25652: contig of 1677 bp in length
25653 25752: gap of unknown length
25753 25697: contig of 845 bp in length
26598 26597: gap of unknown length
26698 26697: gap of unknown length
28740 28740: contig of 2043 bp in length
28741 28840: gap of unknown length
31617 31617: contig of 2777 bp in length
31618 31617: gap of unknown length
31718 33600: contig of 1883 bp in length
33601 33700: gap of unknown length
33701 36667: contig of 2967 bp in length
36668 36767: gap of unknown length
36768 39712: contig of 2945 bp in length
39713 39812: gap of unknown length
39813 42489: contig of 2677 bp in length
42490 42589: gap of unknown length
42590 45916: contig of 3327 bp in length
45917 46016: gap of unknown length
46017 49063: contig of 3047 bp in length
49064 49163: gap of unknown length
49164 52344: contig of 3081 bp in length
52345 52344: gap of unknown length
52345 53907: contig of 3563 bp in length
53907 55007: gap of unknown length
55008 59274: contig of 3267 bp in length
59275 59374: gap of unknown length
59375 61819: contig of 2445 bp in length
61820 61819: gap of unknown length
61920 66981: contig of 5062 bp in length
66982 67081: gap of unknown length
67082 72120: contig of 5039 bp in length
72121 72220: gap of unknown length
72221 75654: contig of 3434 bp in length
75655 75754: gap of unknown length

	*	75755	81137: contig of 5383 bp in length
	*	81138	81237: gap of unknown length
	*	81238	85715: contig of 4478 bp in length
	*	85716	85815: gap of unknown length
	*	85816	91795: contig of 5980 bp in length
	*	91796	91895: gap of unknown length
	*	91896	100931: contig of 9036 bp in length
	*	100932	101031: gap of unknown length
	*	101032	102653: contig of 1622 bp in length.
FEATURES			Location/Qualifiers
SOURCE		1..102653	
		/organism="Lotus corniculatus var. japonicus"	
		/mol_type="genomic DNA"	
		/variety="japonicus"	
		/db_xref="taxon:34305"	
		/chromosome="3"	
		/clone="LjT37F23"	
		/clone_lib="LjT library"	
		/note="TAC clone:TW0506, synonym:Lotus japonicus"	
gap		772..871	
		/estimated_length=unknown	
gap		1868..1967	
		/estimated_length=unknown	
gap		2909..3008	
		/estimated_length=unknown	
gap		3973..4072	
		/estimated_length=unknown	
gap		5829..5928	
		/estimated_length=unknown	
gap		7171..7270	
		/estimated_length=unknown	
gap		8999..9098	
		/estimated_length=unknown	
gap		10096..10195	
		/estimated_length=unknown	
gap		11820..11919	
		/estimated_length=unknown	
gap		13526..13625	
		/estimated_length=unknown	
gap		15096..15195	
		/estimated_length=unknown	
gap		17492..17591	
		/estimated_length=unknown	
gap		19659..19758	
		/estimated_length=unknown	
gap		21398..21497	
		/estimated_length=unknown	
gap		23876..23975	
		/estimated_length=unknown	
gap		25653..25752	
		/estimated_length=unknown	
gap		26598..26697	
		/estimated_length=unknown	
gap		28741..28840	
		/estimated_length=unknown	
gap		31618..31717	
		/estimated_length=unknown	
gap		33601..33700	
		/estimated_length=unknown	
gap		36668..36767	
		/estimated_length=unknown	
gap		39713..39812	
		/estimated_length=unknown	
gap		42490..42589	
		/estimated_length=unknown	
gap		45917..46016	
		/estimated_length=unknown	
gap		49064..49163	
		/estimated_length=unknown	
gap		52245..52344	
		/estimated_length=unknown	
gap		55908..56007	
		/estimated_length=unknown	

gap	59275..59374
	/estimated_length=unknown
gap	61820..61919
	/estimated_length=unknown
gap	66982..67081
	/estimated_length=unknown
gap	72121..72220
	/estimated_length=unknown
gap	75655..75754
	/estimated_length=unknown
gap	81138..81237
	/estimated_length=unknown
gap	85716..85815
	/estimated_length=unknown
gap	91796..91895
	/estimated_length=unknown
gap	100932..101031
	/estimated_length=unknown

ORIGIN

Query Match 12.9%; Score 103.8; DB 14; Length
Best Local Similarity 53.3%; Pred. No. 4.6e-20;
Matches 281; Conservative 0; Mismatches 222; Indels

QY	174	CCCCGAGGTCCTCAGATTATGTGCAAGAACTTTGCGGCTTCTCGGACG	
DB	14335	CTCCAGAGAGTCCCATGTGTGTCATCAACAACTGTGATTTTGTGGAAG	
QY	234	TGGGGTTGTGCTCGAAGTGTACCGAGAGACATGATGCAACGAAGAT	
DB	14275	TGAACATGTGTTCTAAGTGCCAAAGACATGTGCTGAACAGAGCA	
QY	294	CTGAGCAAGCCATCAGGCTGCTCAGGGACATCTGCCACAGCTGCTGC	
DB	14215	CAGCACTCTCATCGAAATATTTATGAATGATATATCAAGCAGCATTTGA	
QY	354	CCGCTCTGTATCATGAGACCAAGCTCACATGCGAGTTGAGAGAAAT	
DB	14155	CTGAACCTGTGTGTGCTGCCA-----CCGTGGAATATCCCACTTAT	
QY	414	ATCAATCTTCCAGCTATTCACAAAGACCTGTTACCCCGCTGCAGCTGC	
DB	14104	CTAAACATATCTCTTTCACAACTGTATTTGGTTTGTCTGTGTTCAAG	
QY	474	TGAAGTCTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATCGAGTGGG	
DB	14044	AGGAAAGCCAAAGATGTCT-----AAAGTTGTAG-----	
QY	534	AGCGTTGTGATTGACAGAGATTTAAGTGTGCTGTGCAACCTCTACTG	
DB	13999	AACGGTTGTGTTGACAGGGTTTAACTGTGATGTGTCATCTTTTCG	
QY	594	GGTACTCGACAAACACACTTGACATATGACTTCAAAAGCCGACAGGCA	
DB	13939	GCTACTCAGACAAACACCATTTGCCATTCGATTACCGACCCGTCGACG	
QY	654	CGAAGCTAATCTCTTGTGCTGCGCCGAGAGGTTCCTAATTTTGA 7	
DB	13879	CTAAGCAAAACCCGTGTGTGAAGGCTGAAGACTTGATTAATCTTAA 1	

RESULT 7	AF139499	1031 bp	mrna	linear	1999
LOCUS	AF139499				
DEFINITION	Prunus armeniaca p85RF mRNA, complete cds.				
ACCESSION	AF139499				
VERSION	AF139499.1	GI:5031280			
KEYWORDS					
SOURCE					
ORGANISM	Prunus armeniaca (apricot)				
	Prunus armeniaca				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta				ta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core				ns;
	rosidae; euroside I; Rosales; Rosaceae; Amygdaloidea				

REFERENCE 1 (bases 1 to 1031)
AUTHORS Mbegwe-A-Mbegwe, D. and File-Lycaon, B.R.
TITLE Molecular cloning and nucleotide sequence of a putative protein
JOURNAL From apricot (Prunus armeniaca var. Bergeron)
UNPUBLISHED
2 (bases 1 to 1031)
REFERENCE Mbegwe-A-Mbegwe, D. and File-Lycaon, B.R.
AUTHORS Direct Submission
TITLE Submitted (31-MAR-1999) INRA, Domaine de Duclos, Petit-Bourg 97170,
JOURNAL France
FEATURES
source Location/Qualifiers
1..1031
/organism="Prunus armeniaca"
/mol_type="mRNA"
/varietal="Bergeron"
/db_xref="taxon:36596"
/cissue_type="fruit mesocarp and endocarp"
/dev_stage="ripe"
1..1031
/gene="p85RF"
113..634
/gene="p85RF"
/note="similar to Phaseolus vulgaris PVPR3 protein"
/codon_start=1
/product="unknown"
/protein_id="ABD38146.1"
/db_xref="GI:5031281"
/translation="MEHETGQPHPEBILCVNNCGFSGSVATRNMSKCHDKMLK
SEQAKLAASFENIVGTSNSNGNEPVVAQVDAHVEPTISLOPSFSGSGG
SGEAREGPKRCGTCKNKRVLGTFNCRGHLFCVAHRSDDKDCPYDHTAADIYAK
ANPVKADKLEKI"
ORIGIN
Query Match 12.7%; Score 102.2; DB 15; Length 1031;
Best Local Similarity 67.8%; Pred. No. 1.2e-19;
Matches 143; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 509 CCCAATCCATGGGGATCTTTCAGAGAGCGTGTGATTCAGAGATTAAAGTGTGGCTGT 568
DB 443 CCAAAACGTTGGGAACTTGCACAAAGAGAGTTGATTAAAGGGTTCAATTGTGGTGT 502
QY 569 GCGAAGCTTCACTGAGCTTATCATCGGTACTCGAGCAACACATTCGACATATGACTAC 628
DB 503 GGTACACCTTTTCTGTCAGTATCATGTTTTCAGACAAACATGACTGCCCTTATGATTAC 562
QY 629 AAAGCCGAGGAGGAGAGAGCGATTGCGAAAGCTAACTCTTGTGTCGTCGCCAGAGGTT 688
DB 563 CACACTGCTGACGCGAGTGTGATGCTAAAGCCACCTGTTAAAGGCTGATAAGCTT 622
QY 689 GTCAAGTTTATGAGCATCCGTTAACTTT 719
DB 623 GAAAAAATCTAAGCCTGATGTGTGAAGTTT 653
RESULT 8
LOCUS BT016379 1020 bp mRNA linear PLN 27-OCT-2004
DEFINITION Zea mays clone Contig212 mRNA sequence.
ACCESSION BT016379
VERSION BT016379.1 GI:54651160
KEYWORDS FLI_CDNA.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 1020)
La1,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Mayer,K.F.X.,
Larkin,B., Becraft,P. and Messing,J.
Characterization of the maize endosperm transcriptome and its
comparison to the rice genome
Unpublished
2 (bases 1 to 1020)
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE

AUTHORS La1,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Ma
Larkin,B., Becraft,P. and Messing,J.
JOURNAL Direct Submission
TITLE Submitted (27-OCT-2004) Waksman Institute, Rutgers
UNPUBLISHED Frelinghuysen Rd, Piscataway, NJ 08854, USA
FEATURES
source Location/Qualifiers
1..1020
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone="Contig212"
ORIGIN
Query Match 12.7%; Score 101.8; DB 15; Length
Best Local Similarity 53.5%; Pred. No. 1.6e-19;
Matches 301; Conservative 0; Mismatches 232; Indels
QY 156 AGGAGACGACCTCGAGGCGCCCTGAGGGTCCAGTTATGTGCAAGACCT 215
DB 117 AGGAGCTGGGTGCGAGGCGCCCGAGGAGCCCATCTCTGCAATATTA 176
QY 216 TCGGACGCCAAGCTACATAGGGGTTGTGCTCGAAGTGTACCGAGAGAC 275
DB 177 TCGGACGCCCGCGGACCATGAACATGTGCTCTAATGTCACAAAGAGAT 236
QY 276 CGAAGATACCGGCTTTAAGCTGAGCAAGCCACTCAGGCTGCTCAGCGAC 335
DB 237 AGGATCAGGCCCAAGCTGAGCTGC-----CTCTCTATGAC 282
QY 336 CTGCTGCTGTTCAAGCCCCCGCTCTCTATGATGAGACCAAGCTCATAC 395
DB 283 CGGACGACCGCGT-----CATGAGCGGTTGTTGCT 328
QY 336 GAACATGATTTGCGCGCATCATCTTCAGCTATCAACAGACCTGGT 455
DB 329 GTAGTAGCTGTTGCCCAAGTGAAGTGAACAAATGAAC--GTGACGCA 386
QY 456 CAGTCCCTTCAGGCAATGAATCTCTTATGCAAGCTCCCTCTAGACC 515
DB 387 TTGCCGGAACCAAGGAGGGGTGGCGGAGATCTCCAAAGGGGAGAGGT 446
QY 516 GATGCGGATCTTTCAGAGAGGGTGTGATTGACAGAGATTAAAGTGTG 575
DB 447 GGTGAGCGCTTCAGAGAGAGGGTGGCTCAGGAGATTCAACTGCCC 506
QY 576 TCTACTGCGCTTTCATGCTGTAATCGACAAACACATCTTGACATATGA 635
DB 507 TGTACTGCGCGCTCCACCGCTACTCCGACAAAGACGACTGCAAGTTGCA 566
QY 636 CAGGCGAGAGAGCGATTGCGAAAGCTAACTCTTGTGTCGCGCGAGAA 695
DB 567 CTGCGAGGAGCGCCATTGCCAAGGCTAATCCGGTGTGAAGGACAGACAN 626
QY 696 TTGATGAGCATCCGTTAACTTT 718
DB 627 TCTAGGGGGGTTCCCTAAGGTT 649
RESULT 9
LOCUS AY15607 1162 bp mRNA linear -2004
DEFINITION Zea mays putative zinc finger protein ZmZF mRNA, cc
ACCESSION AY15607
VERSION AY15607.1 GI:41350258
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; i
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 1162)
Zhang,Z.-X., Tang,W.-H., Tao,Y.-S. and Zheng,Y.-L.
cDNA microarray analysis of early response in maize
Unpublished
2 (bases 1 to 1162)
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE

roots to submerging stress

JOURNAL

Unpublished

2 (bases 1 to 1162)

Zhang,Z.-X., Tang,W.-H., Tao,Y.-S. and Zheng,Y.-L.

AUTHORS

JOURNAL

Submitted (01-JAN-2004) Maize Group of National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University of China, Wuhan, Hubei 430070, China

location/Qualifiers

FEATURES

source

1..1162

/organism="Zea mays"

/mol_type="mRNA"

/culivar="m17"

/db_xref="taxon:4577"

/note="author: Zea mays L."

106..807

/codon_start=1

/product="putative zinc finger protein Zn2f"

/protein_id="AA00453.1"

/db_xref="GI:41350259"

/translation="MVSSPDLKIKWGRGASASRSPSPPLIYLIFPSPFQSS

ERNRKGRLIRKEMEKAKQAPBGPILICINCGFFGSATNMKSKHKENTIK

ODQALVASSIDIVGSDVAVPFGASNTVAVAVQLKKNVQPADVAPVEGV

AAISRGKVGPRCSACRKRVLGTGFNCRGNTLYCALHRYSPDKDKEDYRTAADAI

AKANPVVADKLDKI"

ORIGIN

Query Match 12.5%; Score 100.6; DB 15; Length 1162;

Best Local Similarity 52.2%; Pred. No. 3,8e-19;

Matches 294; Conservative 0; Mismatches 239; Indels 30; Gaps 2;

156 AGGAGACGACCTCGAGAGCCCTGAGGGTCCAGTTATGCGAAGAACCTTGGGCTTCT 215

293 AGGAGGCTGGGTGCGAGGCCCGGAGGAGCCCATCTCTGCAATCAATTAAGTGGCTTCT 352

216 TCGGACGCGCAAGCTACATGAGGTTGTGCTCGAAGTGTCTACGAGAAGCATCATGCAAG 275

353 TCGGACGCGCGGCGACATGAACATGTGCTCGAAGTGCACAAAGAGATGATAGAAAGC 412

276 CGAAGATGACGGCTTTAGCTGAGCAAGCCATCAGGCTGCTCAGGCGACATCTGCCACAG 335

413 AGGATCAGGCGCAAGGTGGCTGC-----CTCCTCTATCGACAGCATGATGA 458

336 CTGCTGCTGTTCAGGCCCCCGCTCTGTATCATGAGCAAGTCAATGCGAGTTGAGA 395

459 CGGACGCGACGCGTATGAGACCGGCTGTGCTGCGACAAACGGTAGTAGTGTTC 518

396 GAAACATGATTTGCGCGCATCAATCTTCCAGCTATCAACAAAGACCTGTTACCCCGCTG 455

519 CCAAGTTAGTTGCAAAAAGAGGTGAGC-----AGCCGCTGATG 562

456 CAGCTGCCCTCAGGACGTGAAGTCTCTATGCGACGCTCCCTCTAGACCCGACCAATC 515

553 TTGCCGAACCCCAACAGAGGGGGTGGCGCGATCTCCAAAGGGGGAAGTAGGGCGAAAC 622

516 GATGCGGATCTTTCAGGAAGCGTTGATGATGACAGATTTAAGTGTGCTGCGCAAC 575

623 GGTGACGCGCTGCGAGAAAGGTTGACTTAAGGATTCAATGCGCGGTGTGGAACT 682

636 TGTACTGCGCACTCCACCGCTACTCGAACAAGACGACTGCAAGTTCCGACTCCGGAAG 742

636 CAGGCGAGAACGATGCGAAAGCTAATCTCTTGTCTGTGCGCGAAGAGGTTGCAAGT 695

743 CTGCCAGGAGACCCATTGCGAAGGCTAATCCGATGTTGAAGCAGACAAAGCTCGACAAGA 802

696 TTTGATGAGCATCGTTAAGCTT 718

803 TCTAGGGGGGTTCCCTACGTT 825

Regult 10

BT017601

LOCUS

Zea mays clone EL01N0431H03.c mRNA sequence.

Accession

BT017601

Version

BT017601.1

Keywords

FLI CDNA.

Source

Zea mays

ORGANISM

Eukaryote; Viridiplantae; Streptophyta; Embryophyta

Spermatophyta; Magnoliophyta; Liliopsida; Poales; P

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 941)

Lai,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Ma

Larking,B., Berafai,P. and Messing,J.

Characterization of the maize endosperm transcriptpro

comparison to the rice genome

Unpublished

2 (bases 1 to 941)

Lai,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Ma

Larking,B., Berafai,P. and Messing,J.

Submitted (27-OCT-2004) Waksman Institute, Rutgers

University, Piscataway, NJ 08854, USA

location/Qualifiers

1..941

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="taxon:4577"

/clone="EL01N0431H03.c"

ORIGIN

Query Match 12.1%; Score 97; DB 15; Length 941

Best Local Similarity 52.9%; Pred. No. 4,9e-18;

Matches 298; Conservative 0; Mismatches 235; Indels

156 AGGAGACGACCTCGAGAGCCCTGAGGGTCCAGTTATGCGAAGAACCT 215

123 AGGAGGCTGGGTGCGAGGCCCGGAGGAGCCCATCTCTGATCAGGGG 182

216 TCGGACGCGCAAGCTACATGAGGTTGTGCTCGAAGTGTCTACCGAGAGC 275

183 TCGGACGCGCGCGACCATGAACTGTCTTAAGTGTGCAAGAGAT 242

276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGAC 335

243 AGGATCAGGCCAAGTGGCTC-----CTCCTCTATCGAC 288

336 CTGCTGCTGTTCAGGCCCGCTCTGTATCATGAGACCAAGCTCAGATG 395

289 CGGACGCGAGCGCGT-----CATGGAGCGGTTGTTGCTG 334

396 GAAACATGATTTGCGCGCATCAATCTTCCAGCTATCAACAAAGACTGAT 455

335 GAGTAGTGTGTTGCCCAAGTGAAGTTGCAACAAATGAAC--GTGACAGA 392

456 CAGTGCCTTCAGGACGATGAAGTCTCTATGCGAGCTCCCTTAGACC 515

453 GGTGACGCGCTGCGAGAAAGGTTGGGCTCACGGAATTCAGTCCG 512

576 TCTACTGCGCTTTACATGCGTACTCGAACAAACACTTGCACATATGA 635

513 TGTACTGGGCGCTCCACGCGTACTCCGACAAAGCAAGCTGCAAGTTCA 572

636 CAGGCGAGAACGATGCGAAAGCTAATCTCTTGTGTCGCGCGAGAA 695

573 CTGCCAGGAGCGCCATTGCGAAGGCTAATCCGATGTTGAAGCAGACAA 632

696 TTTGATGAGCATCGTTAAGCTT 718

Db	633	TCTAGGGGggggtTCCtCAcCGTT	655
RESULT 11			
LOCUS	AR352087	283 bp	DNA
DEFINITION	Sequence 7446 from patent US 6476212.		linear
ACCESSION	AR352087		PAT 20-DEC-2002
VERSION	AR352087.1	GI:27299961	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 283)		
AUTHORS	Lajudl,R.V., Ito,L.Y. and Sherman,B.K.		
TITLE	Polynucleotides and polypeptides derived from corn ear		
JOURNAL	Patent: US 6476212-A 7446 05-NOV-2002;		
FEATURES	Incyte Genomics, Inc.; Palo Alto, CA		
SOURCE	Location/Qualifiers		
	1..283		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	12.0%; Score 96.8; DB 6; Length 283;		
Beech Local Similarity	67.2%; Pred. No. 5.4e-18;		
Matches 137; Conservative	0; Mismatches 67; Indels 0; Gaps 0;		
Oy	509	CCCAATCATGCATCGGATCTTTGGCAGGAAGCGTGTTGGATTGACAGGATTTAAGTCGCCGTG	568
Db	2	CCAAACCGGTGCGCAACTGTAGAAGCGTGTGGGTTGACGGGTTTTAACTGCCGATGC	61
Oy	569	GGAACAACCTCTACTGCGCTTTACATCGGTACTCGAACAAACACACTTGCACATATGACTAC	628
Db	62	GGGAACACGTACTGTTGATGACGCCGTACTCCGACAAACACGATCGCACTTGGACATAT	121
Oy	629	AAAGCCGAGGAGGACAGAAAGCGATTGCCAAAGCTAATCTCTTGTGTCGTCGCGGAGAGGTT	688
Db	122	CGAATCGAGGTAGGGGACGCTATGCCCCAAGGCCAATCCAGTGTGAAGGCGGAGAAAGCTT	181
Oy	689	GTCAGATTGTTGATGAGCATCCGTT	712
Db	182	GACAAAGATCTAGGCGGCGGCAATT	205
RESULT 12			
LOCUS	BC059673	2022 bp	mRNA linear
DEFINITION	Danio rerio zinc finger, A20 domain containing 2, mRNA (cDNA clone		VRT 20-OCT-2004
ACCESSION	MGC:73363 IMAGE:4144947), complete cds.		
VERSION	BC059673		
KEYWORDS	BC059673.1 GI:37589767		
SOURCE	MGC.		
ORGANISM	Danio rerio (zebrafish)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Actinopterygii; Neopterygii; Teleostei; Osteichthyes;		
	Cypriniformes; Cyprinidae; Danio.		
	1 (bases 1 to 2022)		
	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,		
	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,		
	Altschul,S.F., Zeeberg,B., Bluetow,K.H., Schaefer,C.F., Bhat,N.K.,		
	Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F.,		
	Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,		
	Stapleton,M., Soares,M.B., Donald,M.F., Casavant,T.L.,		
	Scheetz,T.E., Brownstein,M.J., Udell,T.B., Toshiyuki,S.,		
	Carninci,P., Prange,C., Rhee,S.S., Loquellano,N.A., Peters,G.J.,		
	Abrahamson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,		
	McMernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,		
	Woley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huylk,S.W.,		
	Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,		
	Raney,J., Helton,E., Kettelman,W., Madan,A., Rodriguez,S.,		
	Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,		

TITLE	Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green
JOURNAL	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,
PUBMED	Butterfield,Y.S., Krzywinski,M.I., Skalka,U., Smal
REFERENCE	Scherech,A., Schein,J.E., Jones,S.J. and Marra,M.A.
AUTHORS	Generation and initial analysis of more than 15,000
TITLE	human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903
PUBMED	12477932
REFERENCE	2 (bases 1 to 2022)
AUTHORS	Director MGC Project.
TITLE	Direct Submission
JOURNAL	Submitted (01-OCT-2003) National Institutes of Health
PUBMED	Gene Collection (MGC), Cancer Genomics Office, Nat'l
REFERENCE	Institute, 31 Center Drive, Room 11A03, Bethesda, P
AUTHORS	USA
TITLE	NIH-MGC Project URL: http://mgc.ncl.nih.gov
JOURNAL	Contact: MGC help desk
PUBMED	Email: cgapbs-remail.nih.gov
REFERENCE	Tissue Procurement: Susan E. Brockerhoff, Universit
AUTHORS	cDNA Library Preparation: Susan E. Brockerhoff, Uni
TITLE	Washington
JOURNAL	cDNA library Arrayed by: The I.M.A.G.E. Consortium
PUBMED	DNA Sequencing by: Sequencing Group at the Stanford
REFERENCE	Center, Stanford University School of Medicine, Sta
AUTHORS	Web site: http://www.sbgc.stanford.edu
TITLE	Contact: (Dickson,Mark) mcdgapaxil.stanford.edu
JOURNAL	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez,
PUBMED	R. M.
FEATURES	Clone distribution: MGC clone distribution informat
SOURCE	through the I.M.A.G.E. Consortium/LINL at: http://j
source	Series: IRAK Plate: 136 Row: c Column: 9
	This clone was selected for full length sequencing
	passed the following selection criteria: matched m
	Location/Qualifiers
	1..2022
	/organism="Danio rerio"
	/mol_type="mRNA"
	/strain="wild-type"
	/db_xref="taxon:7955"
	/clone="MGC:73363 IMAGE:4144947"
	/tissue_type="Eye, adult retina, 1-2 year
	/clone_lib="Zebrafish adult retina cDNA"
	/lab_host="MDH10B"
	/note="vector: pbluescript SK-"
	1..2022
	/gene="za20d2"
	/note="synonyms: id:fbid2892, wu:fb16d02, v
	fb16d02, fc67h07, 2892"
	/db_xref="GeneID:406312"
	/db_xref="ZFIN:ZDB-GENE-040426-1979"
	129..770
	/gene="za20d2"
	/codon_start=1
	/product="zinc finger, A20 domain containi
	/protein_id="AAHS9673.1"
	/db_xref="GI:37589768"
	/db_xref="GeneID:406312"
	/db_xref="ZFIN:ZDB-GENE-040426-1979"
	/translation="MAQETNQSPVPIICTGCGGYGNPRTNG
	DRSEMPLAGSPSAEAIAIORLEASINKALPAPESTETME
	SISREBKALPKATVEVPITOPTSSPSIPVAQANDAKSP
	VGLTGPRCRGNLFEGIHRSYDKNCTVDYKAEMAIIRKEN
ORIGIN	
Query Match	11.8%; Score 94.6; DB 5; Length 20
Best Local Similarity	66.3%; Pred. No. 2.8e-17;
Matches 136; Conservative 0; Mismatches 69; Indels	0;
Oy	480 CCTTATCGAGCCTCCCTTAGACCCGACCCCATTCATGGGATCTTC
Db	550 CCAAGAGCCGAGACTCTCCCAACCAAGAAAATAAGATGCTTTACCT
	539
	609

Qy 626 TACAAAGCCGACAGGAGAGAGATTCGGAAGCTATCTCTTGTCTGCGCCGAGAG 685
Db 456 TATCGAAGGCTGCTCAAGATGCTATTCGAAAGCCACCAAGTTGTTAAGGCTGAGAAAG 515
Qy 686 GTTGTCAGTTTGAAG 702
Db 516 CTTGACAAATATGAG 532

Search completed: December 8, 2005, 09:09:44
Job time : 4261 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2005, 06:02:18 ; Search time 552 Seconds
(without alignment)
9707.270 Million cell updates/sec

Title: US-10-716-089-12

Perfect score: 804
Sequence: 1 atcccgagcaccagtcgccg.....agtcgcatcagaagctcgcc 804

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	804	100.0	804	6	ABST0622 Transcrip
2	804	100.0	804	6	AA895832 CDNA enco
3	776.4	96.6	1531	6	ABST0614 Transcrip
4	776.4	96.6	1531	6	AA895824 Partial c
5	103.4	12.9	1003	13	ADK47821 Plant ful
6	103	12.8	873	13	ADK53121 Plant ful
7	103	12.8	945	13	ADK51406 Plant ful
8	102.6	12.8	932	13	ADK51034 Plant ful
9	102.2	12.7	1032	13	ADK52834 Plant ful
10	101.8	12.7	776	13	ADK47307 Plant ful
11	101.8	12.7	821	13	ADK47342 Plant ful
12	101.8	12.7	826	13	ADK46752 Plant ful
13	101.8	12.7	841	13	ADK50951 Plant ful
14	101.8	12.7	846	13	ADK50987 Plant ful
15	101.8	12.7	853	13	ADK54045 Plant ful
16	101.8	12.7	863	13	ADK52432 Plant ful
17	101.8	12.7	863	13	ADK53800 Plant ful
18	101.8	12.7	868	13	ADK33018 Plant ful
19	101.8	12.7	868	13	ADK33018 Plant ful

ALIGNMENTS

20	101.8	12.7	868	13	ADK54028	Adx	ful
21	101.8	12.7	872	13	ADK51702	Adx	ful
22	101.8	12.7	876	13	ADK52889	Adx	ful
23	101.8	12.7	886	13	ADK12369	Adx	ful
24	101.8	12.7	894	13	ADK48800	Adx	ful
25	101.8	12.7	898	13	ADK61184	Adx	ful
26	101.8	12.7	899	13	ADK50721	Adx	ful
27	101.8	12.7	902	13	ADK52776	Adx	ful
28	101.8	12.7	904	13	ADK47138	Adx	ful
29	101.8	12.7	918	13	ADK61202	Adx	ful
30	101.8	12.7	920	13	ADK47807	Adx	ful
31	101.8	12.7	924	13	ADK60812	Adx	ful
32	101.8	12.7	926	13	ADK52877	Adx	ful
33	101.8	12.7	928	13	ADK52869	Adx	ful
34	101.8	12.7	942	13	ADK55778	Adx	ful
35	101.8	12.7	943	13	ADK5613	Adx	ful
36	101.8	12.7	948	13	ADK10810	Adx	ful
37	101.8	12.7	953	13	ADK53073	Adx	ful
38	101.8	12.7	956	13	ADK53049	Adx	ful
39	101.8	12.7	959	13	ADK47752	Adx	ful
40	101.8	12.7	976	13	ADK47829	Adx	ful
41	101.8	12.7	978	13	ADK32914	Adx	ful
42	101.8	12.7	985	13	ADK47720	Adx	ful
43	101.8	12.7	987	13	ADK62462	Adx	ful
44	101.8	12.7	989	13	ADK48942	Adx	ful
45	101.8	12.7	990	13	ADK46608	Adx	ful

RESULT 1	ABST0622	standard; CDNA, 804 BP.
ID	ABST0622;	
AC	ABST0622;	
DT	28-NOV-2002	(first entry)
XX		
DE	Transcription factor stress-related protein (TFSRP), cDNA	
XX		
KW	Transcription factor stress-related protein; TFSP; MYB-1	
KW	AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;	
KW	ZF-4 protein; ZF-5 protein; CAAT-Box binding factor prot;	
KW	Sigma factor like protein; SFU-1; transgenic; plant; droug	
KW	environmental stress; Physcomitrella patens; high salinity	
KW	ciliate; fungus; gene; ss.	
XX		
OS	Physcomitrella patens.	
XX		
PN	US2002102695-A1.	
XX		
PD	01-AUG-2002.	
XX		
PF	06-APR-2001; 2001US-00828303.	
XX		
PR	07-APR-2000; 2000US-0196001P.	
XX		
PA	(SILV/) SILVA O D C E.	
PA	(BOHN/) BOHNERT H J.	
PA	(THIE/) THIELEN N V.	
PA	(CHEN/) CHEN R.	
PI	Silva ODE, Bohnert HJ, Thiele N, Chen R;	
DR		
DR	WPI; 2002-690614/74.	
XX		
DR	P-PSDB; ABG93809.	
XX		
PT	Novel transcription factor stress-related protein such as	
PT	protein, zinc-finger protein, or CAAT-Box binding factor i	
XX	useful for modifying stress tolerance of plant.	
XX		
PS	Claim 3; Fig 2D; 104pp; English.	

xx The invention relates to an isolated transcription factor stress-related protein (TFSPR) (I) selected from AP2 similar-2 protein (AP2-2), zinc-finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein, ZF-5 protein, MYB-1 protein, CMT-Box binding factor protein-3 (CMT-3), Sigma factor like protein (SFL-1) and their orthologues. Also described is: (1) an isolated TFSPR coding nucleic acid (II) coding for (I); (2) a transgenic plant cell (III) transformed by (II), where the expression of (II) in the plant cell results in increased tolerance to an environmental stress as compared to a wild-type variety of the plant cell; (3) a transgenic plant (IV) comprising (III); (4) a seed (V) produced by (IV), where the seed is true breeding for an increased tolerance to environmental stress as compared to a wild-type variety of the plant cell; (5) an agricultural product (VI) produced by (IV) or (V); and (6) an isolated recombinant expression vector (VII) comprising (II), where expression of (VII) in a host cell results in increased tolerance to environmental stress as compared to a wild-type variety of the host cell. (II) is useful for modulating (increasing or decreasing) stress tolerance of a plant (transgenic or not transgenic), by modifying the expression of TFSPR in the plant. The plant is transformed with a promoter that directs expression of the TFSPR. The promoter is tissue specific and is developmentally regulated. TFSPR expression is modified by administration of an antisense molecule that inhibits expression of TFSPR. (VII) is useful for producing a transgenic plant containing (II), where expression of the nucleic acid in the plant results in increased tolerance to environmental stress as compared to a wild-type variety of the plant, by transforming a plant cell with (VII) comprising the nucleic acid, generating from the plant cell a transgenic plant with an increased tolerance to environmental stress as compared to a wild-type variety of the plant. (I) is useful for conferring stress tolerance such as drought, cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful for identifying Physcomitrella patens and related organisms, for mapping of genomes of organisms related to P. patens, for identifying and localising P. patens sequences of interest, for evolutionary studies, for determining TFSPR regions required for function, for modulating TFSPR activity, for modulating metabolism of one or more compounds, and for modulating stress resistance. (I) is useful for reducing stress tolerance response plants or more particularly, in the transcription of a protein involved in a stress tolerance response in a P. patens plant. (II) is useful for transforming plants and thus inducing tolerance to stresses such as drought, high salinity and cold, for identifying the presence of P. patens or a related organism in a mixed population of microorganisms, serve as marker for specific regions of the genome for mapping the genome and for the functional studies of P. patens proteins. (II) is useful as reference points for mapping the moss genomes, or of genomes of related organisms, for evolutionary and protein structural studies, for generating knockout mutation in the genomes of various organisms such as bacteria, mammalian cells, yeast cells and plant cells, useful for evaluating their ability or capacity to tolerate various stress conditions and the effect on the phenotype and/or genotype of the mutation. (II) is useful as marker for the construction of a genomic map in related mosses. (I) or (II) is useful for generating algae, ciliates, plants, fungi or other microorganisms expressing mutated TFSPR nucleic acid and protein molecules for improving stress tolerance. ABS70611-ABS70681 represent P. patens TFSPR coding sequences and PCR primers of the invention.

SQ Sequence 804 BP, 186 A; 214 C; 205 G; 199 T; 0 U; 0 Other;

Query Match 100.0%; Score 804; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 2.3e-250;
Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCCGGGACCAAGTCCGCTTAGTGTGTCTATTAGTGTGGTTGCAAGTCTGAAGC 60
|||
DB 1 ATCCGGGACCAAGTCCGCTTAGTGTGTCTATTAGTGTGGTTGCAAGTCTGAAGC 60
|||

QY 61 CTTGAGCGAGATTTCAGAGATTTCATATAGCTTCTGATTGGAAGTATACCTTAT 120
|||
DB 61 CTTGAGCGAGATTTCAGAGATTTCATATAGCTTCTGATTGGAAGTATACCTTAT 120
|||

QY 121 TAGTCTGTTAAAGATGGCCAGCGGTGTCTCAGAGAGCAACTCGACGCCCTGA 180
|||

DB 121 TAGTCTGTTAAAGATGGCCAGCGGTGTCTCAGAGAGCAACTCG 180
|||
QY 181 GGGTCCAGTTATGTGCAAGAACCTTTGGCGGCTTCTTGGGCGCCAAAGT 240
|||
DB 181 GGGTCCAGTTATGTGCAAGAACCTTTGGCGGCTTCTTGGGCGCCAAAGT 240
|||

QY 241 GTGCTCGAAGTGTACCGAGAGACAGTCATCAGACGCAAGATGACGGCT 300
|||
DB 241 GTGCTCGAAGTGTACCGAGAGACAGTCATCAGACGCAAGATGACGGCT 300
|||

QY 301 AGCCACTCAGGCTGCTCAGCGGACATCTGCCACAGCTGCTGCTTACG 360
|||
DB 301 AGCCACTCAGGCTGCTCAGCGGACATCTGCCACAGCTGCTGCTTACG 360
|||

QY 361 TGTACATGAGCCAAAGCTCAGATCGAGGTTGAGAGAACATGATTTGTC 420
|||
DB 361 TGTACATGAGCCAAAGCTCAGATCGAGGTTGAGAGAACATGATTTGTC 420
|||

QY 421 TTCCAGCTATCAGACAGACCTGTTACCCCGCTGACAGTCCCTCAG 480
|||
DB 421 TTCCAGCTATCAGACAGACCTGTTACCCCGCTGACAGTCCCTCAG 480
|||

QY 481 CTCTATCGACGCTCCCTTACAGCCGAGCCCAATCGATGCGGATCTTGC 540
|||
DB 481 CTCTATCGACGCTCCCTTACAGCCGAGCCCAATCGATGCGGATCTTGC 540
|||

QY 541 TGTGATGACAGGATTTAAGTGTGCTGTGGCAACCTCTACGCGCTTTA 600
|||
DB 541 TGTGATGACAGGATTTAAGTGTGCTGTGGCAACCTCTACGCGCTTTA 600
|||

QY 601 GGAACAAACACCTTGACATATGACTACAAAGCCGACGAGGACGAGAGCG 660
|||
DB 601 GGAACAAACACCTTGACATATGACTACAAAGCCGACGAGGACGAGAGCG 660
|||

QY 661 TAACTCTTGTGTGCTGGCGGAGAAAGTTGTCAAGTTTGTATGACATTC 720
|||
DB 661 TAACTCTTGTGTGCTGGCGGAGAAAGTTGTCAAGTTTGTATGACATTC 720
|||

QY 721 CTGCGCAGATTAGGCTTCAATGATGATGATCAATCAATCTTCTTC 780
|||
DB 721 CTGCGCAGATTAGGCTTCAATGATGATGATCAATCAATCTTCTTC 780
|||

QY 781 AGCGAGTCGATCAAGAGCTCGCC 804
|||
DB 781 AGCGAGTCGATCAAGAGCTCGCC 804
|||

RESULT 2
AAS95832
ID AAS95832 standard; cDNA; 804 BP.
XX
AC AAS95832;
XX
AC 26-FEB-2002 (first entry)
XX
DE cDNA encoding zinc finger-4 (ZF-4).
XX
XX Transcription factor stress-related protein; AP2 similar-2
KM AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;
KM MYB-1 MYB-1 protein; CMT-box binding factor protein-3; CMT
KM Sigma factor like protein; SFL-1; environmental stress; mo.
XX 88.
OS Physcomitrella patens.
XX
PN MO200177311-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US011393.
XX
PR 07-APR-2000; 2000US-0196001P.

er;

XX (BADI) BASF PLANT SCI GMBH.
 XX Da Costa Silva O, Bohnert HJ, Van Thielén N, Chen R;
 XX WPI; 2002-049146/06.
 XX P-PSDB; AANU1800.
 XX
 XX New polypeptide, useful for identification of Physcomitrella patens, and
 XX for modulating stress resistance of a plant, comprises an isolated
 XX transcription factor stress-related protein derived from Physcomitrella
 XX patens.
 XX
 XX Claim 14; Fig 2D; 129pp; English.
 XX
 XX The invention relates to an isolated transcription factor stress-related
 XX protein (TFSRP) (i) selected from an AP2 similar-2 protein (APS-2), a
 XX zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 protein, ZF-5
 XX protein, a MYB-1 protein (MYB-1), a CAAT-box binding factor protein-3
 XX (CABF-3), a sigma factor like protein (SFL-1) and its orthologues. Also
 XX described is an isolated TFSRP coding nucleic acid (ii) which codes for
 XX (i); and an isolated recombinant expression vector (iii) comprising (ii),
 XX where expression of (iii) in a host cell results in increased tolerance
 XX to environmental stress as compared to a wild type variety of the host
 XX cell. (i) or (ii) is useful as markers for specific regions of the TFSRP
 XX genome. (i) or (ii) is useful for identification of Physcomitrella patens
 XX and related organisms, mapping of genomes of organisms related to P.
 XX patens, identification and localization of P. patens sequence of
 XX interest, evolutionary studies, determination of TFSRP regions required
 XX for function, modulation of a TFSRP activity, modulation of the
 XX metabolism of one or more cell functions, modulation of the transmembrane
 XX transport of one or more compounds and modulation of stress resistance.
 XX (iii) is useful for generating probes and primers for identifying and/or
 XX cloning TFSRP homologues in other cell types and other organisms, as well
 XX as TFSRP homologues from mosses and related species, and for evolutionary
 XX CC and protein structural studies. AAG595821-AAG595891 represent P. patens
 XX TFSRP coding sequences and PCR primers of the invention
 XX
 XX Sequence 804 BP; 186 A; 214 C; 205 G; 199 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 804; DB 6; Length 804;
 XX Best Local Similarity 100.0%; Pred. No. 2,3e-250;
 XX Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 ATCCGGGACCAAGCTCCGCTTATGTTGTCATTAAGTGTGTTGACGCTGAAGC 60
 XX 1 ATCCGGGACCAAGCTCCGCTTATGTTGTCATTAAGTGTGTTGACGCTGAAGC 60
 XX
 XX 61 CTGAGCGAGATTGCGAGATTCTTCATACGCTTCTGATTAGAAAGTACACCTTAT 120
 XX 61 CTGAGCGAGATTGCGAGATTCTTCATACGCTTCTGATTAGAAAGTACACCTTAT 120
 XX
 XX 121 TAGTGTGTTAAAGATGCGACGAGCGTGTCTCAGAGACGACCTCGCAGCCCTCTGA 180
 XX 121 TAGTGTGTTAAAGATGCGACGAGCGTGTCTCAGAGACGACCTCGCAGCCCTCTGA 180
 XX
 XX 181 GGGTCCAGTTATGTGTGCAAGAACCTTTGCGGCTTCTCGGACCAAGCTACATGCGGATT 240
 XX 181 GGGTCCAGTTATGTGTGCAAGAACCTTTGCGGCTTCTCGGACCAAGCTACATGCGGATT 240
 XX
 XX 241 GTGCTCGAAGTGTCTACCGAGACAGTGTGCAAGGAAAGATGCGCTTATGCTGAGAGA 300
 XX 241 GTGCTCGAAGTGTCTACCGAGACAGTGTGCAAGGAAAGATGCGCTTATGCTGAGAGA 300
 XX
 XX 301 AGCCACTCAGGCTGCTCAGGCGACATCTGCCACAGCTGTGCTTTACGCCCCGCTCC 360
 XX 301 AGCCACTCAGGCTGCTCAGGCGACATCTGCCACAGCTGTGCTTTACGCCCCGCTCC 360
 XX
 XX 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAGAACAAATGATTTGCGGCATCAATC 420
 XX 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAGAACAAATGATTTGCGGCATCAATC 420
 XX
 XX 421 TTCGAGCTATCAACAAGACCTGTTATCCCGCTGTGAGTGCCTTCAGGCAAGTGAATC 480
 XX 421 TTCGAGCTATCAACAAGACCTGTTATCCCGCTGTGAGTGCCTTCAGGCAAGTGAATC 480

DB 421 TTCGAGCTATCAACAAGACCTGTTATCCCGCTGTGAGTGCCTTCAGGCAAGTGAATC 480
 QY 481 CTCTATTCGAGCTCCCTCTAGACCCGAGCCCAATGCATGCGATCTTGC 540
 DB 481 CTCTATTCGAGCTCCCTCTAGACCCGAGCCCAATGCATGCGATCTTGC 540
 QY 541 TCGATTGACAGATTTAAGTGTGCTGTGCGCAACCTTACCTGCGCTTTT 600
 DB 541 TCGATTGACAGATTTAAGTGTGCTGTGCGCAACCTTACCTGCGCTTTT 600
 QY 601 GGACAAACACACTTGACATATGACTACAAAGCCGACAGGCGAGAACGC 660
 DB 601 GGACAAACACACTTGACATATGACTACAAAGCCGACAGGCGAGAACGC 660
 QY 661 TAATCTCTTGTGTCGCGGAGAGAGTTGTCAAGTTTGTATGATGAGATCC 720
 DB 661 TAATCTCTTGTGTCGCGGAGAGAGTTGTCAAGTTTGTATGATGAGATCC 720
 QY 721 CTGCGGAGATTTAGGCTTCAATCAATTGAGTAACTCTACATCTTTCTTC 780
 DB 721 CTGCGGAGATTTAGGCTTCAATCAATTGAGTAACTCTACATCTTTCTTC 780
 QY 781 AGCGAGTGCATCAAGAGCTGCGC 804
 DB 781 AGCGAGTGCATCAAGAGCTGCGC 804
 XX
 XX RESULT 3
 XX ABS70614
 XX ID ABS70614 standard; cDNA; 1531 BP.
 XX
 XX ABS70614;
 XX
 XX 28-NOV-2002 (first entry)
 XX
 XX Transcription factor stress-related protein (TFSRP), parti
 XX
 XX Transcription factor stress-related protein; TFSRP; MYB-1
 XX AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;
 XX ZF-4 protein; ZF-5 protein; CAAT-Box binding factor protei
 XX Sigma factor like protein; SFL-1; transgenic; plant; droug
 XX environmental stress; Physcomitrella patens; high salinity
 XX ciliate; fungus; gene; ss.
 XX
 XX Physcomitrella patens.
 XX
 XX US2002102695-A1.
 XX
 XX 01-AUG-2002.
 XX
 XX 06-APR-2001; 2001US-00828303.
 XX
 XX 07-APR-2000; 2000US-0196001P.
 XX
 XX (SILV/) SILVA O D C E.
 XX (BOHN/) BOHNERT H J.
 XX (THIE/) THIELEN N V.
 XX (CHEN/) CHEN R.
 XX
 XX Silva ODE, Bohnert HJ, Thielén NV, Chen R;
 XX WPI; 2002-690614/74.
 XX
 XX Novel transcription factor stress-related protein such as
 XX protein, zinc-finger protein, or CAAT-Box binding factor p
 XX useful for modifying stress tolerance of plant.
 XX
 XX Example 5; Fig 1D; 104pp; English.
 XX
 XX The invention relates to an isolated transcription factor
 XX protein (TFSRP) (i) selected from AP2 similar-2 protein (i/
 XX finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein;

MYB-1 protein, CAAT-Box binding factor protein-3 (CABF-3), Sigma factor like protein (SPF-1) and their orthologues. Also described is: (1) an isolated TFSRP coding nucleic acid (II) coding for (I); (2) a transgenic plant cell (III) transformed by (II), where the expression of (II) in the plant cell results in increased tolerance to an environmental stress as compared to a wild-type variety of the plant cell; (3) a transgenic plant (IV) comprising (III); (4) a seed (V) produced by (IV), where the seed is true breeding for an increased tolerance to environmental stress as compared to a wild-type variety of the plant cell; (5) an agricultural product (VI) produced by (IV) or (V); and (6) an isolated recombinant expression vector (VII) comprising (II), where expression of (VII) in a host cell results in increased tolerance to environmental stress as compared to a wild-type variety of the host cell. (II) is useful for modifying (increasing or decreasing) stress tolerance of a plant (transgenic or not transgenic), by modifying the expression of TFSRP in the plant. The plant is transformed with a promoter that directs expression of the TFSRP. The promoter is tissue specific and is developmentally regulated. TFSRP expression is modified by administration of an antisense molecule that inhibits expression of TFSRP. (VII) is useful for producing a transgenic plant containing (II), where expression of the nucleic acid in the plant results in increased tolerance to environmental stress as compared to a wild-type variety of the plant, by transforming a plant cell with (VII) comprising the nucleic acid, generating from the plant cell a transgenic plant with an increased tolerance to environmental stress as compared to a wild-type variety of the plant. (I) is useful for conferring stress tolerance such as drought, cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful for identifying phycomitrella patens and related organisms, for mapping of genomes of organisms related to P.patens, for identifying and localizing P.patens sequences of interest, for evolutionary studies, for determining TFSRP regions required for function, for modulating TFSRP activity, for modulating metabolism of one or more cell functions, for modulating transmembrane transport of one or more compounds, and for modulating stress resistance. (I) is useful for reducing stress tolerance response plants or more particularly, in the transcription of a protein involved in a stress tolerance response in a P.patens plant. (II) is useful for transforming plants and thus inducing tolerance to stresses such as drought, high salinity and cold, for identifying the presence of P.patens or a related organism in a mixed population of microorganism, serve as marker for specific regions of the genome for mapping the genome and for the functional studies of P.patens proteins. (II) is useful as reference points for mapping the moss genomes, or of genomes of related organisms, for evolutionary and protein structural studies, for generating knockout mutation in the genomes of various organisms such as bacteria, mammalian cells, yeast cells and plant cells, useful for evaluating their ability or capacity to tolerate various stress conditions and the effect on the phenotype and/or genotype of the mutation. (II) is useful as marker for the construction of a genomic map in related mosses. (I) or (II) is useful for generating algae, ciliates, plants, fungi or other microorganisms expressing mutated TFSRP nucleic acid and protein molecules for improving stress tolerance. AB570611-AB570681 represent P. patens TFSRP coding sequences and PCR primers of the invention

Sequence 1531 BP, 313 A, 393 C, 393 G, 426 T, 0 U, 0 Other;

Query Match 96.6%; Score 776.4; DB 6; Length 1531; Beest Local Similarity 99.7%; Pred. No. 3.1e-241;

Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

8 GCACCAAGTCCCGCTTAGTGTGTGTCATTAAGTGTGGTGAAGCTTGAGC 67
Db GCACCAAGTCCCGCTTAGTGTGTGTCATTAAGTGTGGTGAAGCTTGAGC 588
68 GAGATTGAGGATTTTCTCATACGCTTGTGATTAAGAAAGATACACCTTATTAAGCTG 127
Db GAGATTGAGGATTTTCTCATACGCTTGTGATTAAGAAAGATACACCTTATTAAGCTG 648
128 TTAAGAGTGGCCACCGAGGCTGTGCTCAGSAGAGACCTCGCAGGCCCTGAGGCTCA 187
Db TTAAGAGTGGCCACCGAGGCTGTGCTCAGSAGAGACCTCGCAGGCCCTGAGGCTCA 708
188 GTTATGTGCAGAACCTTTCGCGCTTCTTCGCGACCAAGCTAACATGAGGTTGTGCTCG 247

|||||
Db GTTATGTGCAGAACCTTTCGCGCTTCTTCGCGACCAAGCTAACATG 768
709 GTTATGTGCAGAACCTTTCGCGCTTCTTCGCGACCAAGCTAACATG 768
QY AAGGCTTACCGAGAGACAGTATGCAAGCCGAAGATGACGCTTTAGCTG 307
Db AAGGCTTACCGAGAGACAGTATGCAAGCCGAAGATGACGCTTTAGCTG 827
QY CAGGCTGTCCAGGCGACACTGTCGCCACAGCTGCTGTTCAGCCCCCG 367
Db CAGGCTGTCCAGGCGACACTGTCGCCACAGCTGCTGTTCAGCCCCCG 887
QY GAGACCAAGCTTCACATGCGAGGTTGAGAGAACATGATTTGCGCATC 427
Db GAGACCAAGCTTCACATGCGAGGTTGAGAGAACATGATTTGCGCATC 947
QY TATCAACAAGCTGTGTTACCCCGCTGACAGTCCCTCAGCAGTGA 487
Db TATCAACAAGCTGTGTTACCCCGCTGACAGTCCCTCAGCAGTGA 1007
QY GAGCTCCCTTACAGCCGAGCCCAATGATGCGGATTTGACAGAAC 547
Db GAGCTCCCTTACAGCCGAGCCCAATGATGCGGATTTGACAGAAC 1067
QY ACAGGATTTAAGTGTGCTGTGGAACCTTACTGCGCTTTACATCGGT 607
Db ACAGGATTTAAGTGTGCTGTGGAACCTTACTGCGCTTTACATCGGT 1127
QY CACACTTCACATATGATACAAAGCCGACAGGCGAGAGAGATTTGCA 667
Db CACACTTCACATATGATACAAAGCCGACAGGCGAGAGAGATTTGCA 1187
QY CACACTTCACATATGATACAAAGCCGACAGGCGAGAGAGATTTGCA 1187
668 CTGTGCGTGGCCGAGAGAGTGTCAAGTTTGATGAGCATCCGTTAAC 727
Db CTGTGCGTGGCCGAGAGAGTGTCAAGTTTGATGAGCATCCGTTAAC 1247
QY CGATTAGGCTTCATATGATGATGATGATGATGATGATGATGATGATG 787
Db CGATTAGGCTTCATATGATGATGATGATGATGATGATGATGATGATG 1307
1248 CGATTAGGCTTCATATGATGATGATGATGATGATGATGATGATGATG 1307
QY CGCATCAAGA 797
Db CGCATCAAGA 1317

RESULT 4

AAS95824 standard; cDNA; 1531 BP.

AAS95824;

26-FEB-2002 (first entry)

Partial cDNA encoding Zinc finger-4 (ZF-4).

Transcription factor stress-related protein; AP2 similar-2

AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;

MYB-1 MYB-1 protein; CAAT-box binding factor protein-3; Cai

Sigma factor like protein; SPF-1; environmental stress; mo

Physcomitrella patens.

MO200177311-A2.

18-OCT-2001.

06-APR-2001; 2001MO-US011393.

07-APR-2000; 2000US-0196001P.

(BADI) BASF PLANT SCI GMBH.

Da Costa Silva O, Bohnert HJ, Van Thielén N, Chen R;

et;

CC available in electronic form from the US patent office at
 CC ftp://seeddata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 1003 BP; 211 A; 266 C; 294 G; 232 T; 0 U; 0 Other;

Query Match 12.9%; Score 103.4; DB 13; Length 1003;
 Best Local Similarity 53.6%; Pred. No. 2.1e-22;
 Matches 302; Conservative 0; Mismatches 231; Indels 30; Gaps 3;

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTCATGTAATGCAAGAACCTTTGGCCTTCT 215
 DB 122 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCATCTCTGATCAATAGTGGCTTCT 181
 QY 216 TCGGCAGCCAGACTACATGAGGCTGCTCGAAGTGTCAACGAGAGACATGCAAG 275
 DB 182 TCGGCAGGCGCGGCGACATGAAACATGTGCTCCAGTGCACAAAGAGATGATTAAGAAAGC 241
 QY 276 CGAAGATGACGCGCTTGAAGTGAAGCAAGCACTCAGGCTGCTGAGGCGACATGCGACAG 335
 DB 222 AGGATCAGGCGCAAGCTGGCTGC-----CTTCCTTATGACAGATGTGAA 287
 QY 336 CTGCTGCTGTTAGCGCCCCCTCTCTGTAATGATGAGACCAAGCTCAATGCGAGTTGAGA 395
 DB 288 CGGAGCGACGCGCT-----CATGAGCGCGTGTGTTGCTGGAGCAAGACG 333
 QY 336 GAAACATGATTTGCGCGCATCATCTTCAGCTATTCACAAAGACCTGGTTACCCCGCTG 455
 DB 334 GTATGATGCTGTTGCCCAAGTTGATGTAACAAATGAAC--GTGACAGACGCCGCTGATG 391
 QY 456 CAGCTGCCCTCAGGAGTGAAGTCCCTATGCGAGCTCCCTGTAACCCGAGCCCAATC 515
 DB 332 TTGCGGACCCAGGAGGCGGCTGGCGGCGCATCTCAAGGGGGGAAAGTAGGGCCGAAC 451
 QY 516 GATGCGGATCTTTGACAGAACGCTGTGATTTGACAGATTTAACTGCTGTGCAAC 575
 DB 452 GGTGAGGCGCTGACAGAAAGAGGTTGACTTACGGGATTCAACTGCCGGTGTGGAACT 511
 QY 576 TCTACTGCGCTTATCATGCTGTACTGGACAAACACATCTTGCAATATGACTTCAAGCCG 635
 DB 512 TGTACTGGGCACTTCCACCGCTACTCCGACAGACAGACTGCAAGTTCGACTACCCGACTG 571
 QY 636 CAGGCGAGGAGCGATTTGCGAAAGCTAATCTTGTGTGCGCGAGGAAGTGTGCAGT 695
 DB 572 CTGCGAGGAGCGCATTTGCCAAGCTAATCCGCTGTGTGAAGGACAGACAGCTCGACAAAG 631
 QY 696 TTTGATGAGCATCGCTTAAGCTT 718
 DB 632 TCTAGGGGGGGTCTTACGGTT 654

RESULT 6
 ADX53121
 ID ADX53121 standard; cDNA; 873 BP.
 XX
 AC ADX53121;
 XX 21-APR-2005 (first entry)
 DT
 XX Plant, full length insert polynucleotide seqid 27861.

XX plant proectant; plant growth regulant; gene therapy; pla
 KW recombinant DNA construct; physical array; plant breeding
 KW cold tolerance; heat tolerance; drought tolerance; herbici
 KW extreme osmotic condition; pathogen tolerance; pest tolera
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth
 KW yield; plant growth; plant development; seed oil; protein
 KW protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/J.) LIU J.

XX (ZHOU/J.) ZHOU Y.

XX (KOVA/J.) KOVALIC D K.

XX (SCRE/J.) SCREEN S E.

XX (TABAJ/) TABASKA J E.

XX (CAO/Y.) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

XX WPI; 2004-180133/17.

XX Claim 1; SEQ ID NO 27861; 15pp; English.

CC The invention describes a recombinant DNA construct compri
 CC polynucleotide consisting of a sequence encoding an amino
 CC available in electronic form from the US patent office at
 CC ftp://seeddata.uspto.gov/sequence.html?docid:2004034888. The
 CC of the invention are also useful in physical arrays of mol
 CC plant breeding markers. The recombinant DNA construct is u
 CC improving plant tolerance to cold, heat, drought, herbicid
 CC osmotic conditions, pathogens or pests, for manipulating g
 CC plant cells by modification of the cell cycle pathway, for
 CC increased resistance to plant disease, for producing galac
 CC lignin or plant growth regulators, for increasing the rate
 CC recombination in plants, for improving yield by modificati
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use
 CC or by providing improved plant growth and development unde
 CC stress condition or for modifying seed oil or protein yield
 CC content. This sequence represents a plant full length inse
 CC polynucleotide that can be used in the recombinant DNA con
 CC invention.

XX Sequence 873 BP; 180 A; 251 C; 264 G; 178 T; 0 U; 0 Other;

Query Match 12.8%; Score 103; DB 13; Length 87.
 Best Local Similarity 53.2%; Pred. No. 2.7e-22;
 Matches 307; Conservative 0; Mismatches 240; Indels

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTCATGTAATGCAAGAACTT 215
 DB 144 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCATCTCTGATCAATTAAC 203
 QY 216 TCGGCAGCCAGACTACATGAGGCTGCTCGAAGTGTCAACGAGAGAC 275
 DB 204 TCGGCAGGCGCGGCGACATGAAACATGTGCTTAAAGTGCACAAAGAGAT 263
 QY 276 CGAAGATGACGCGCTTGAAGTGAAGCAAGCACTCAGGCTGCTCAGGCGAC 335

Db 264 AGGATCAGGCCAGAGCTGCTGC-----CTCTCTATGAGACGATCTGTGA 309
Qy 336 CTGCTGCTGTTGACGCCCGCTCTCTGATGATGAGCAAGCTCATGCGAGTTGAGA 395
Db 310 CGGACGCGACGCCGT-----CATGAGCCGGTTGTTGCTGCGACGACACG 355
Qy 396 GAAACATGATTTGTGCGCGCATCTTCCAGCTATTAACAAGACCTGGTTACCCCGCTG 455
Db 356 GTATGAGCTGTGGCCCAAGTCAGATTGCAAAACATGAMAC--GTGCAGCAGCCCGCTGATG 413
Qy 456 CAGCTGCCCTCAGGAGGATGAAGTCTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATC 515
Db 414 TTGCGGACCCAGCAGAGGGGGTGGCGGCATCTCCAAAGGGGGAAGTAGGGCCGACCC 473
Qy 516 GATGCGGATCTTTCAGAGAAAGCGTTGTGATTAAGATTTAAGTGTGCGTGGCAACC 575
Db 474 GGTGACGCGCTGACAGAGAGGGGTGGGCTCACGGGATTTCAACTGCCGTGGGAACT 533
Qy 576 TCTACTGCGCTTTACATCGGTACTCTGAGAACAAACACACTTGACATATGACTACAAAGCCG 635
Db 534 TGTACTGCGCGCTCCACCGCTACTCTGACAGACAGACGACGAACTTCACTCCGGAATG 593
Qy 636 CAGGCGAGGAAGCGATTGCGAAAGCTAATCTTGTGCGGCCGAGAAAGTTGTCAAGT 695
Db 594 CTGCCAGGAGCGCCATCTGCCAAGGCTAATCCGGTGTGAAGGACAGACACTCGACACAGA 653
Qy 696 TTTGATGAGCATCCGTTAAGCTTTCTGCGCAGCAT 732
Db 654 TCTAGGGGGGGTTCCTCAGCGTTGTGTGTCAGGAAGATT 690

RESULT 7

ADXS1406
ID ADXS1406 standard; cDNA; 945 BP.
XX AC
XX ADXS1406;
DT 21-APR-2005 (first entry)
DE Plant full length insert polynucleotide seqid 26146.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS unidentified.
XX
PN US2004034888-A1.
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIU/J) LIU J.
PA (ZHOU/J) ZHOU Y.
PA (KOVA/J) KOVALIC D K.
PA (SCRE/J) SCREEN S E.
PA (TABAR/J) TABASKA J E.
PA (CAO/Y) CAO Y.
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
DR WPI; 2004-180133/17.
XX

PT New recombinant DNA construct, useful for improving plant
PT cold, heat, drought, herbicides, extreme osmotic condition
PT pests, for conferring increased resistance to plant disease
PT improving yield.
PS Claim 1, SEQ ID NO 26146, 15bp; English.

CC The invention describes a recombinant DNA construct compri
CC polynucleotide consisting of a sequence encoding an amino
CC available in electronic form from the US patent office at
CC ftp.segdata.nebto.gov/sequence.html?docid:2004034888. The
CC of the invention are also useful in physical arrays of mo
CC plant breeding markers. The recombinant DNA construct is
CC improving plant tolerance to cold, heat, drought, herbicide
CC osmotic conditions, pathogens or pests, for manipulating
CC plant cells by modification of the cell cycle pathway, fo
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yiel
CC content. This sequence represents a plant full length inse
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

XX
SQ Sequence 945 BP; 207 A; 245 C; 256 G; 237 T; 0 U; 0 Other;
Query Match 12.8%; Score 103; DB 13; Length 94
Best Local Similarity 52.3%; Pred. No. 2.8e-22;
Matches 284; Conservative 0; Mismatches 220; Indels

Qy 156 AGGAGACGACCTCGAGGCGCCCTGAGGGTCCAGTTATGCAAGACCT 215
Db 113 AGGAGACTGAGATGCGACGACCTGAGGAGACCATCTTTGACATTA. 172
Qy 216 TCGGAGCCAGGCTACATGAGGGTGTGCTGGAATGCTACCGAAGAC 275
Db 173 TCGGACGCGAGCTACATGAACATGTGCTCCAAATGCGACAAAGAGAT 232
Qy 276 CGAAGATGACGCGCTTAACTGAGCAAGCACTCAGGCTGTCAGCGAC 335
Db 233 AGGACGAGGCCAAGTGTGCTG----- 253
Qy 336 CTGCTGCTGTTACGCCCCCGCTCTGTAATGATGACCAAGCTCATG 395
Db 254 CTTCTTATGACAGCATGCGTCAATGGAACGATGCTGTCATGGAAC 313
Qy 396 GAAACATGATTTGTGCGCATTAATTTCCAGCTATCAACAAAGCTGT 455
Db 314 GCACACAGTGTGCTGCTGCTGCTCAATCGAGTTGCAAAACATGAACG 373
Qy 456 CAGTGCCTCAGGAGGATGAAGTCTCTATGCGAGCTCCCTCTAGAC 515
Db 374 ATGTTGCTGAGACTAGCGAGGGGGCGCGGTGATCTCCAAAGGGAAGT 433
Qy 516 GATGCGGATCTTTCAGAGAAAGCGTGTGATTAAGATTTAAGTGTG 575
Db 434 GGTGACGACCTTTCAGAGAAAGGGTGGATTGACATTAAGATTCAAGTCC 493
Qy 576 TCTACTGCGCTTTACATGCGTATCTGGAACAAACACTTGACATATGA 635
Db 494 TGTACTGTGCACTGACCGCTACTCCGACAGACGACGACTGCAAGTTGCA 553
Qy 636 CAGGCGAGGAAGCGATTGCGAAAGCTAATCTCTGTGTCGCGCAGAA 695
Db 554 CTGCTAGGGATGCCATTGCCAAGGCTAATCCAGTGTGAAGCGGACAA 613
Qy 696 TTT 698
Db 614 TCT 616

RESULT 8
 ADX51034
 ID ADX51034 standard; cDNA, 932 BP.
 XX
 AC
 XX ADX51034;
 DT 21-APR-2005 (first entry)
 XX
 XX Plant full length insert polynucleotide seqid 25774.
 DE
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX
 OS unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 25774; 15pp; English.
 XX
 XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photoassimilates or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SO Sequence 932 BP; 205 A; 236 C; 264 G; 227 T; 0 U; 0 Other;

Oy	82	TTTCTCATACGCTTCTGATTAGAGAAAGTATACACCTTATTATCTGTAA	141
Db	136	TGTGACACTACTATTCTGGATTATGATCAGATTGTTTGGTATTATTCTAGTAT	195
Oy	142	CGAGCGTGTGTCTCAGAGAGACGACCTCCGACGGCCCTGTAGGGTCCAGTTT	201
Db	196	GGAG-----ACAAGAGAGGCTGGGTGCGAGGCCCCCGAGGGAGCCATC	249
Oy	202	CCTTTGCGGCTTTCTTCGGCAGCCCAAGCTAACCTATGGGGTTGTCTCGAAGT	261
Db	250	TAACTGTGGCTTCTTCGGCAGCGCGCCACCACTGAAACAATGTCTCCAGT	309
Oy	262	GACAGTCATGCAAGCGAAGATGACCGGCTTTAGCTGACGCAAGCACACTCAG	321
Db	310	GATGATTAACGAAGCAGAGATCAGGCCCAACTGGCTGC-----	355
Oy	322	GACATCTGCGCAAGCTGCTGCTGTTCAGCCCCCGCTCTCTGTACATGAG	381
Db	356	GACAGCATCGTGAAGCGGACGACGCGT-----CATGGA	401
Oy	382	ATGCGAGGTTGAGAGAACAAATGATGTGTCGCCGATCATATCTTCCAGCTAT	441
Db	402	CTGGCAGCAACACGGTAGTAGCTGTGCCCAAGTTGATGTGCACAAACAT	459
Oy	442	GATTACCCCGCTGCAGCTGCCCTCAGAGCAGTGAAGTCCCTATCGCA	501
Db	460	GCAGCCGCTGATGTGTGCGGAGCCCAAGGAGGGGGTGGGGCGGATCTCC	519
Oy	502	ACCCGAGCCCAATGATGTCGGATCTTTCAGAGAAAGGCTGTGGATTGACA	561
Db	520	GGTAGGGCCGAACCGGTGCACGCTTCGAGGAAGGGTTGATCTTAACG	579
Oy	562	TGCTGTGGCAACCTCTACTGCGCTTTACATCGGTACTGAGCAAAACAC	621
Db	580	CCGGTGTGGAACTTGTACTGTGCACTCCACCGCTACTCTCCAGCAAGCAC	639
Oy	622	TGACTTCAAAAGCCGCAAGGAGGAAGCAATTCGAAAGCTAAATCCTCTT	681
Db	640	CGACTACCGGACTGCTGCGCAAGGAGACGCATTGCGCAAGGCTAATCCGGTG	699
Oy	682	GAGGTGTCAAGTTTGTATGATGAGCATCCGTTAAGCTT 718	
Db	700	CAAGCTGACAAAGATCTTAGGGGGGTTTCCTACGGTT 736	
RESULT 9			
ADX52834			
ID	ADX52834	standard; cDNA; 1032 BP.	
XX	ADX52834;		
AC			
XX	21-APR-2005 (first entry)		
DT			
XX			
DE	Plant full length insert polynucleotide seqid 27574.		
XX			
KM	plant protectant; plant growth regulant; gene therapy; pla		
KM	recombinant DNA construct; physical array; plant breeding		
KM	cold tolerance; heat tolerance; drought tolerance; herbici		
KM	extreme osmotic condition; pathogen tolerance; pest tolera		
KM	growth rate; cell cycle pathway; disease resistance;		
KM	galactomannan production; lignin production; plant growth		
KM	yield; plant growth; plant development; seed oil; protein		
KM	protein content; gene; ss.		
XX			
OS	Unidentified.		
XX			
PM	US2004034888-A1.		
XX			
PD	19-FEB-2004.		
XX			
PF	28-APR-2003; 2003US-00425114.		
XX			
PR	06-MAY-1999; 99US-00304517.		

PR 05-NOV-2001; 2001US-00985678.
 XX (LIU/J) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,
 XX WPI, 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 27574; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 1032 BP; 211 A; 278 C; 304 G; 239 T; 0 U; 0 Other;
 Query Match 12.7%; Score 102.2; DB 13; Length 1032;
 Best Local Similarity 51.3%; Pred. No. 5.3e-22;
 Matches 288; Conservative 0; Mismatches 248; Indels 27; Gaps 1;
 OY 156 AGAGACGACCTCGAGGCGCCCTGAGGGTGCAGTTATGCAAGAACTTTGGCGCTCT 215
 DB 135 AGAAGGCTGGGTGCGAGGCGCCCGAGGGACCATCTCTGATCAATATAGTGGCTTCT 194
 OY 216 TCGGACGCCAAGCTACATGGGGTGTGTCGAAGTGTCAAGAGACATGATGCAAG 275
 DB 195 TCGGACGGCGCGACCATGATGATGTCTTAACTGTGCGCAAGAGATGATGAAGAAC 254
 OY 276 CGAAGATGACGGCTTAACTGAGCAAGCACTCAGAGCTGCTGAGCGCATGTGCACAG 335
 DB 255 AGGA-----TCAGGCCAAGCTGTGCTCTCTTATTTG 287
 OY 336 CTGCTGCTGTACGCCCCCGCTCTGTATCATGAGAACCAAGTCAATGCGAGTTGAGA 395
 DB 288 ACAGCATGTGAACGGACGACGCGCTCATGAGACCGGTTGGTGTGCGACACAGAG 347
 OY 396 GAAACATATATTGTGCGCATCAATCTTCAGCTATCAACAGAACTGGTTACCCCGCTG 455
 DB 348 TAGTAGCTGTGGCCCAAGTGTGCAAAACATGAACGTGACAGACCGCTGTATGTTG 407
 OY 456 CAGCTGCCCTCAGAGCAAGTAAGTCTCTATCGAGCTCCCTCTAGACCCGAGCCCAATC 515
 DB 408 CCGACACCAAGCAGAGGGGGTGGCGCGCATCTCCAAAGGGGGGAAGTAGGGCGGAACC 467
 OY 516 GATGCGGATCTTTGACAGAAAGCTGTTGATGACAGATTAAAGTGTGCTGTGCAACC 575
 DB 468 GGTGACGGCGCTGACAGAAAGAGGTTGGGCTCACGGGATTCAAATGCGGGTGTGGAAC 527

OY 576 TCTACTGGGCTTTACATGCTGCTCGACAAACACACTTGCATATAG, 635
 DB 528 TGTACTGGCGGCTCCACCGCTACTCCGACAAAGACAGACTGCAAGTTGCA, 587
 OY 636 CAGGCGAGGAAGCGATTGCGAAAGCTAAATCTTGTGCGTGGCCGAGAA, 695
 DB 588 CTGCGACGGAGCGCATTTGCGAAAGCTAAATCGGGTGTGGAAGCAGACAA, 647
 OY 696 TTTGATGAGCATCGCTTAAGCTT 718
 DB 648 TCTAGGGGGGTTCCCTTACGGTT 670
 RESULT 10
 ADX47307
 ID ADX47307 standard; cDNA; 736 BP.
 XX
 AC ADX47307;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DB Plant full length insert polynucleotide seqid 22047.
 XX
 KW plant protectant; plant growth regulant; gene therapy; pla
 KW recombinant DNA construct; physical array; plant breeding
 KW cold tolerance; heat tolerance; drought tolerance; herbic
 KW extreme osmotic condition; pathogen tolerance; pest toler
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth
 KW yield; plant growth; plant development; seed oil; protein
 KW protein content; gene; ss.
 XX
 OS Undefined.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 XX
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIU/J) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao
 XX WPI, 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant
 PT cold, heat, drought, herbicides, extreme osmotic condition
 PT pests, for conferring increased resistance to plant diseas
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 22047; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct compri
 CC polynucleotide consisting of a sequence encoding an amino
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The
 CC of the invention are also useful in physical arrays of mol
 CC plant breeding markers. The recombinant DNA construct is u
 CC improving plant tolerance to cold, heat, drought, herbicid
 CC osmotic conditions, pathogens or pests, for manipulating g
 CC plant cells by modification of the cell cycle pathway, for
 CC increased resistance to plant disease, for producing galac
 CC lignin or plant growth regulators, for increasing the rate
 CC recombination in plants, for improving yield by modificati

CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 736 BP; 158 A; 213 C; 225 G; 140 T; 0 U; 0 Other;

Query Match 12.7%; Score 101.8; DB 13; Length 736;
Best Local Similarity 53.5%; Pred. No. 6e-22;
Matches 301; Conservative 0; Mismatches 232; Indels 30; Gaps 3;

QY 156 AGAAGACGACCTCCGAGGCGCCCTGAGGGTCCAGTTATGTGCAAGAACCTTTGGCTTCT 215
DB 114 AGAAGGCTGGTGGCCAGGCCCCCGAGGAGCCCATCTCTCTGATCAATTAATCTGGCTTCT 173
QY 216 TCGGAGCCCAAGCTACATGGGGTGTGCTCGAAGTGTCTACCGAGAGACAGTCATGCAAG 275
DB 174 TCGGAGGCGCGCGACATGAACATGTGCTCCAGTGGCCAGAGAGATGATTAAGAAAGC 233
QY 276 CGAATATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTAGGCGACATCTGCACAG 335
DB 234 AGGATCAGGCCCAAGCTGGCTGC-----CTCTCTATTCGACAGCATGTGAA 279
QY 336 CTGCTGCTGTTTCAAGCCCCCGCTCCGTATACATGAGACCAAGTCAATGCGAGTTGAGA 395
DB 280 CGGACGCCAGCCGCT-----CATGAGCCGCGTTTGTGCTGGCAGCAACAGC 325
QY 396 GAACATATGTTGTGCGCATCATCTTCCAGCTATCAACAGACCTGTGTTACCCCGCTG 455
DB 336 GATGATAGCTGTGGCCAGTTGAGTTGCAAAATGAAC--GTGACAGAGCCCGCTGATG 383
QY 456 CAGCTGCCCTTCAAGGCAATGAGTCTTATGTGCAAGTCCCTCTGTAGACCCGACCAATC 515
DB 384 TTGCCCGAACCCAGCGAGGGGTGGCGGCATCTCCAAAGGGGGAAGGTAGGGCCGAAC 443
QY 516 GATGCGGATCTTGGACGAAGCGGTGTGATGACAGATTTAAGTGTGCTGGCAAC 575
DB 444 GGTGACGCCCTTGGAGAGAGAGGTTGACTTACGGGATTTCACTGCCGTGTGGAACT 503
QY 576 TCTACTGCGCTTATCATCTGCTGCTGCAACACACTTGCATATGACTACAAAGCCG 635
DB 504 TGTACTGGGCACTCCACCGCTACTCGACAAACAGACGAAAGTTGACATCCGGAACG 563
QY 636 CAGGCGAAGAACGCTTGTGCAAAAGCTAATCTCTTGTCTGTGCGCGAAGAGTTTCAAGT 695
DB 564 CTGCCAGGAGCCCATTTGCAAGGCTAATCCGCTGTGAGAGCAGACAGCTCGACAAGA 623
QY 696 TTTGATGAGCATCCGTTAAGCTT 718
DB 624 TCTAGGGGGGTTTCCCTACGCTT 646

RESULT 11

ADXS3784
ID ADXS3784 standard; cDNA; 776 BP.

ADXS3784;

21-APR-2005 (first entry)

Plant full length insert polynucleotide seqid 28524.

XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.

OS unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 05-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/J.) LIU J.

XX (ZHOU/Y.) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABBA/) TABASKA J E.

XX (CAO/Y.) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

XX WPI; 2004-180133/17.

XX Claim 1; SEQ ID NO 28524; 15bp; English.

CC The invention describes a recombinant DNA construct compris
CC polynucleotide consisting of a sequence encoding an amino
CC available in electronic form from the US patent office at
CC ftp:seedata.uspto.gov/sequence.html?docid:2004034888. The i
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yield
CC content. This sequence represents a plant full length inse
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

XX Sequence 776 BP; 162 A; 225 C; 238 G; 151 T; 0 U; 0 Other;

Query Match 12.7%; Score 101.8; DB 13; Length
Best Local Similarity 53.5%; Pred. No. 6.2e-22;
Matches 301; Conservative 0; Mismatches 232; Indels 3;

QY 156 AGAAGACGACCTCCGAGGCGCCCTGAGGGTCCAGTTATGTGCAAGAACCT 215
DB 134 AGAAGGCTGGTGGCCAGGCCCCCGAGGAGCCCATCTCTGATCAATTA 193
QY 216 TCGGAGCCCAAGCTACATGGGGTGTGCTCGAAGTGTCTACCGAGAGC 275
DB 194 TCGGAGCGCGCGACCATGAAACATGTGCTTAAGTGCACACAGAGAT 253
QY 276 CGAATATGACGGCTTTAGCTGAGCAAGCACTCAGGCTGCTCAGGCGAG 335
DB 254 AGGATCAGGCCCAAGCTGCTGC-----CTCTCTATTCGAC 299
QY 336 CTGCTGCTGTTCAAGCCCCCGCTCTGTACATGAGACCAACCTCAGATG 395
DB 300 CGGACGCCAGCCGCT-----CATGAGCCGCGTTGTGCTG 345
QY 396 GAACATATGTTGTGCGCATCATCTTCCAGCTATCAACAAAGACCTGCT 455
DB 346 GTAGTAGCTGTGGCCCAAGTGAAGTTCGAAACAAATGAAC--GTGACGCA 403

QY 456 CAGCTGCCCTCAGGACGTAAGTCTCTATTCGAGCTCCCTCTAGACCCGACCCCAATC 515
DB 404 TTGCCGACCCACGAGGGGGGTGGCGCGATCTCCAAAGGGGGAAGTGAAGCCGAAAC 463
QY 516 GATCGGATCTTTCAGGAAAGCGTTGGATGACAGGATTTAAGTTCGCTGAGGCAAC 575
DB 464 GGTGACGCGCTTCAGGAAAGGATGGGCTCAGCGGATTCACCTGCGGTGGGAACT 523
QY 576 TCTACTGCGCTTTACATCGGTAATCGGACAAACACATTCGACATATGACTACAAAGCCG 635
DB 524 TGTACTGCGCGCTTCACCGCTACTCCGACAAAGCAGCTGCAAGTTGACTACGGAACT 583
QY 636 CAGGCGAAGAAAGGATTTGCAAAAGCTATCTCTTGTTCGTGCGCGGAGAGGTTGCAAGT 695
DB 584 CTGCGAGGAGCGCATTTGCGCAAGCTAATCCGGTGTGAAGGACGCAAGCTCGACAGA 643
QY 696 TTTGATGAGCATCGTTAAGCTT 718
DB 644 TCTAGGGGGGTTCCCTACGGTT 666

RESULT 12

ADK47342

ID ADK47342 strand: cDNA; 821 BP.

AC ADK47342;

DT 21-APR-2005 (first entry)

DE Plant full length insert polynucleotide seqid 22082.

KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
OS Unidentified.
XX US2004034888-A1.
PN 19-FEB-2004.PD 28-APR-2003; 2003US-00425114.
PF 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.

XX (LIU/J) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
PI MPI; 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

XX Claim 1; SEQ ID NO 22082; 15bp; English.

XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at

CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificat
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yield
CC content. This sequence represents a plant full length ins
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

SQ Sequence 821 BP; 171 A; 236 C; 247 G; 167 T; 0 U; 0 Other;

Query Match

Best Local Similarity 12.7%; Score 101.8; DB 13; Length

Matches 301; Conservative 0; Mismatches 232; Indels 3;

QY 156 AGGAGACGACCTCGCAGGCGCCCTGAGGCTCCAGTTATGTGCAAGAACCT
DB 114 AGGAGCTGGGTGCCAGGCGCCCGAGGAGCCCATCTCTGCTCATATA
QY 216 TCGGACGCCAAGCTACATGAGGTTGTGCTGCAAGTCTACCGAGAC
DB 174 TCGGACGCCGCGGACCATGAAATGTGCTCCAAAGTCCCAAGAGAT
QY 276 CGAATGACGCGCTTAACTGAGCAAGCACTCAGGCTGCTCAGGAGAC
DB 234 AGATTCAGGCCAAGCTGAGTGC-----CTCCTCTATGAC
QY 336 CTGCTGCTGTTCAGCGCCCGCTCCTGATCATGAGACCAAGCTCATG
DB 280 CCGACCGACGCGCT-----CATGAGCCGCGTTGTGCTG
QY 396 GAACATGATGTGCGCATCAATCTTCAGCTATCAACAAAGCTGAT
DB 326 GTAGTAGCTGTGCCAAGTTGAGTTGCAACATGAAAC--GTGACAGC
QY 456 CAGCTGCCCTCAGGACGTAAGTCTCTATGCGACGCTCTCTAGAC
DB 384 TTGCCGGAACCCAGCGAGGGGGTGGCGGCGATCTCCAAAGGGGGAAGT
QY 516 GATCGGATCTTGAAGAAAGCGTTGATGACAGGATTTAAGTGC
DB 444 GGTGACGCGCTGCAAGAAAGGATTTGACTTACGGGATTCACATGCC
QY 576 TCTACTGCGCTTTACATCGGTAATCGGACAAACACATTCGACATATG
DB 504 TGTACTGCGCACTTCACCGCTACTCCGACAAAGCAGCTGCAAGTTG
QY 636 CAGGCGAAGAAAGGATTTGCAAAAGCTATCTTGTCTGCGCGAGAA
DB 564 CTGCGAGGAGCGCATTTGCGCAAGCTAATCCGGTGTGAAGCAGACAA
QY 696 TTTGATGAGCATCGTTAAGCTT 718
DB 624 TCTAGGGGGGTTCCCTACGGTT 646

RESULT 13

ADK46752

ID ADK46752 strand: cDNA; 826 BP.

AC ADK46752;

DT 21-APR-2005 (first entry)

DE Plant full length insert polynucleotide seqid 21492.

XX	plant protectant; plant growth regulator; gene therapy; plant;
KM	recombinant DNA construct; physical array; plant breeding marker;
KM	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KM	extreme osmotic condition; pathogen tolerance; pest tolerance;
KM	growth rate; cell cycle pathway; disease resistance;
KM	galactomanan production; lignin production; plant growth regulator;
KM	yield; plant growth; plant development; seed oil; protein yield;
KM	protein content; gene; ss.
XX	Unidentified.
OS	
XX	US2004034888-A1.
XX	
PD	19-FEB-2004.
XX	
PF	28-APR-2003; 2003US-00425114.
XX	
PR	06-MAY-1999; 99US-00304517.
PR	05-NOV-2001; 2001US-00985678.
XX	
PA	(LIU/) LIU J.
PA	(ZHOU/) ZHOU Y.
PA	(KOVA/) KOVALIC D K.
PA	(SCRE/) SCREEN S E.
PA	(TABA/) TABASKA J E.
PA	(CAO/) CAO Y.
XX	
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,
XX	
DR	WPI; 2004-180133/17.
XX	
PT	New recombinant DNA construct, useful for improving plant tolerance to
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT	pests, for conferring increased resistance to plant disease, or for
PT	improving yield.
XX	
PS	Claim 1; SEQ ID NO 21492; 15pp; English.
CC	
CC	The invention describes a recombinant DNA construct comprising a
CC	polynucleotide consisting of a sequence encoding an amino acid sequence
CC	available in electronic form from the US patent office at
CC	ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC	of the invention are also useful in physical arrays of molecules and as
CC	plant breeding markers. The recombinant DNA construct is useful for
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in
CC	plant cells by modification of the cell cycle pathway, for conferring
CC	increased resistance to plant disease, for producing galactomanan,
CC	lignin or plant growth regulators, for increasing the rate of homologous
CC	recombination in plants, for improving yield by modification of
CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC	or by providing improved plant growth and development under at least one
CC	stress condition or for modifying seed oil or protein yield and/or
CC	content. This sequence represents a plant full length insert
CC	polynucleotide that can be used in the recombinant DNA construct of the
CC	invention.
XX	
SQ	Sequence 826 BP; 169 A; 217 C; 252 G; 188 T; 0 U; 0 Other;
XX	
QY	Query Match 12.7%; Score 101.8; DB 13; Length 826;
	Best Local Similarity 53.5%; Pred. No. 6.4e-22;
	Matches 301; Conservative 0; Mismatches 233; Indels 30; Gaps 3
DB	156 AGGAGAGCACTCCGACAGGCCCTTGAGGCTCCAGTTATGCAAGAACTTTGGCGCTTCT 215
	9 AGGAGGCTGGGTGCGCAGGCCCTCCGAGGACCACCTCTCTGCATCAATAACTGTGGCTTCT 68
QY	216 TCGGACGCCAAGCTACATCCATGGGGTTGTGCTCGAAGTCTCTACCGAGAGCACTGTCAG 275
DB	69 TCGGACAGCGCGGACCACTGAACAATGTCTCTAATGTGCACACAGAGAGATGATTAAGAA 128
QY	276 CGAAGATGACGGCTTTAGCTGACGAAGCACTCAGGCTGCTCAGCGCATCTGCGCAG 335

Db	129	AGGATCAGGCGCAAGCTGGGTGC-----CTCCTTATGCAC
Qy	336	CTGCTGCTGTTTACGCCCCCGCTCCTGTTACATGAGCCAAAGTCACATG
Db	175	CGGCAGCGCAACGCCGT-----CATGGAACCGGTTGTGCTG
Qy	336	GAACCATGATTGTGCGCGCATCATCTTCCAGCTATCAACAAAGACTGGT
Db	221	GTAAGTAGCTGTGGCCCAAGTCCAGTTGCACAAATGAAAC--GTGCAGCA
Qy	456	CAGCTGCCCTTCAGGCGAGTAAAGTCTCTATGCGAGCTCCCTCTAGACC
Db	279	TTGCCGGAACCCAGCGAGGGGGGTGGCGCGCATCTCCAAGGGGGAAAGGT
Qy	516	GATGCGGATCTTGCAGGAAGCGTGTGGAATTGACAGATTTTAAGTGTG
Db	339	GGTGACAGCGCTTGCAGGAAGAGGTTGGGCTCACGGGATTCACATGCGC
Qy	576	TCTACTGGCGCTTACATGCGGTACTCGGACAAACACACTTGACATATGA
Db	399	TGTACTGGCGCTTCCACCGCTACTCCGACAAAGCAGACTGCAAGTTGA
Qy	636	CAGGCGAGGAAGCGATTGCGAAAGCTAATCCTTGTGCGCCGAGAA
Db	459	CTGCCAGGGAGCGCATTTGCCAAGGCTAATCCGGTGTGTAAGCAGACA
Qy	696	TTTGATGAGCATCCGTTAAGCTT 718
Db	519	TCTAGGGGGGGTTCCTTACGCTT 541
RESULT 14		
ADX50951		
ID	ADX50951	standard; cDNA, 841 BP.
XX	XX	
AC	ADX50951;	
XX	XX	
DT	21-APR-2005	(first entry)
DE	Plant full length insert polynucleotide seqid 25691.	
XX	XX	
KM	plant protectant; plant growth regulant; gene therapy; plant	
KM	recombinant DNA construct; physical array; plant breeding	
KM	cold tolerance; heat tolerance; drought tolerance; herbicide	
KM	extreme osmotic condition; pathogen tolerance; pest tolera	
KM	growth rate; cell cycle pathway; disease resistance;	
KM	galactomannan production; lignin production; plant growth	
KM	yield; plant growth; plant development; seed oil; protein	
KM	protein content; gene; ss.	
OS	Unidentified.	
XX	XX	
PN	US2004034888-A1.	
XX	XX	
PD	19-FEB-2004.	
XX	XX	
PF	28-APR-2003; 2003US-00425114.	
XX	XX	
PR	06-MAY-1999; 99US-00304517.	
PR	05-NOV-2001; 2001US-00985678.	
XX	XX	
PA	(LITUJ/) LIT J.	
PA	(ZHOU/) ZHOU Y.	
PA	(KOVALIC/) KOVALIC D K.	
PA	(SCRE/) SCREEN S E.	
PA	(TABAR/) TABASKA J E.	
PA	(CAOY/) CAO Y.	
XX	XX	
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao	
XX	XX	
DR	WPI; 2004-180133/17.	
XX	XX	
PT	New recombinant DNA construct, useful for improving plant	

PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

PS Claim 1, SEQ ID NO 25691, 15pp, English.

CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing Galactomannan,
CC liginin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 841 BP; 177 A; 237 C; 252 G; 175 T; 0 U; 0 Other;

Query Match 12.7%; Score 101.8; DB 13; Length 841;

Best Local Similarity 53.5%; Pred. No. 6.4e-22;

Matches 301; Conservative 0; Mismatches 232; Indels 30; Gaps 3;

156 AGGAGACGACCTCGCAGGCGCCCTGAGGGTTCGAGTTATGTGCAAGAACCTTGGCGCTTCT 215
113 AGGAGGCTGGGGCCAGGCGCCCGAGGAGCCATCTCTGCAATCATATACCTGTGGCTTCT 172
216 TCGGCGAGCCAGCTACCATGAGGGTGTGCTCGAAGTGTCTACGAGAGACAGTCATGCAAG 275
173 TCGGCGAGCGCGCGAGCAGTGAACATGTGCTCCAGAGTCCCAAGAGAGATGATACGAAAC 232
276 CGAAGATGACGGCTTATGCTGAGCAAGCAGCTCAGGCTGCTCAGGAGCATGTGCGACAG 335
233 AGGATCAGGCGCAGCTGCTGC-----CTCCTCTATGCAAGCATGCTGTA 278
336 CTGCTGCTGTTAGCCCCCGCTCCTGTCATGAGACCAAGCTCACATGCGAGTTGAGA 395
279 CGGCGAGCGAGCGCT-----CATGAGCGCGGTTGCTGCGAGCAACAG 324
396 GAACATGATGTGCGCGCATCATCTTCAGGCTATCAACAAAGACCTGGTTACCCCGCTG 455
325 GTAGTAGCTGTGGCCCAAGTTGAGTTGCAAAATGAAC--GTGACAGCAGCCCGCTGATG 382
456 CAGCTGCCCTCAGGAGAGTGAAGTCTCTATGCGAGCTCCTCTAGAACCCGAGCCCATC 515
383 TTGCGCGAGCCAGCGAGGGGGGTGGGGCGATCTCCAAAGGGGGGAAGTAAAGGCGGAAAC 442
516 GATGCGGATCTTGCAGAAAGCTGTGATTTGAAGATTTAAGTGTGCTGTGCAAC 575
443 GGTGAGGCGCTCGAGAGAGAGGTTGAGCTTACGGGATTCAGCTGCCGTGTGGAACT 502
576 TCTACTGCGCTTAACTCGGTACTGCGAACAACACACTTGCAATATGACTCAAAAGCGG 635
503 TGTACTGCGCATCTCAACCGCTACTCGAGAACAGAGACTGCAAGTTGCGACTACCGGACTG 562
636 CAGGCGAGAGACGATTTGCGAAAGTAACTCTTGTGCTGCGAGCGAGAGAGTTGTCAAGT 695
563 CTGCGCAGGAGCGCATTTGCGCAAGGCTAATCGGTGTGAAAGCAGACAGCTCGACAGA 622
636 TTTGATGAGCATCCGTTAAGCTT 718
623 TCTAGGGGGGGTTCCTTACGGTT 645

RESULT 15

ADX50987
ID ADX50987 standard; cDNA, 846 BP.

AC ADX50987;

XX 21-APR-2005 (first entry)

DE Plant full length insert polynucleotide seqid 25727.

XX plant protectant; plant growth regulant; gene therapy; pla
KW recombinant DNA construct; physical array; plant breeding;
KW cold tolerance; heat tolerance; drought tolerance; herbici
KW extreme osmotic condition; pathogen tolerance; pest tolera
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; liginin production; plant growth
KW yield; plant growth; plant development; seed oil; protein
XX protein content; gene; ss.

OS Unidentified.

PN US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

PA (LILU/) LIU J.

PA (ZHOY/) ZHOU Y.

PA (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.

PA (TABAS/) TABASKA J E.

PA (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

DR WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant

PT cold, heat, drought, herbicides, extreme osmotic condition

PT pests, for conferring increased resistance to plant disease

PT improving yield.

XX Claim 1, SEQ ID NO 25727, 15pp; English.

XX The invention describes a recombinant DNA construct compri
CC polynucleotide consisting of a sequence encoding an amino
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galec
CC liginin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yiel
CC content. This sequence represents a plant full length inse
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

XX Sequence 846 BP; 178 A; 241 C; 252 G; 175 T; 0 U; 0 Other;

Query Match 12.7%; Score 101.8; DB 13; Length

Best Local Similarity 53.5%; Pred. No. 6.4e-22;

Matches 301; Conservative 0; Mismatches 232; Indels

156 AGGAGACGACCTCGCAGGCGCCCTGAGGGTTCGAGTTATGTGCAAGAACCT

```
Db 118 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCATCCTCTGCACTCAATACTGTGGCTTCT 177
Oy 216 TCGGCAGCCAAAGCTAACCATGGGGTTGTGCTGAAAGTGTACCGAGAGACAGTCATGCAAG 275
Db 178 TCGGCAGCGCGCGCACATGAAACATGTGCTCCAAAGTGCACAAAGAGATGATTAACGAAGC 237
Oy 276 CGAAGATGACCGGCTTACTGAGCAAGCACTCAGGCTGCTCAGGCGACATCTGCCACAG 335
Db 238 AGGATCAGGCCCAGAGCTGGCTGC-----CTCCTTATCGAAGCATGTGAA 283
Oy 336 CTGCTGCTGTTTCAGCCCCCGCTCCTGTACATGAGACCAAGCTCACATGCGAGTTGAGA 395
Db 284 CGGCAAGCAACGCGT-----CANGAGCCGGTTGTGTGCTGGCAGCAACG 329
Oy 396 GAACATGATTTGTGCGCATTCATCTTCAGCTATCAACAAAGACCTGTTACCCCGCTG 455
Db 330 GTAGTAGCTGTGCCCCAAGTTGAGTTCAAACATGAAC--GTGCAGCAGCCCGCTGATG 387
Oy 456 CAGCTGCCCCCTCAGGCAATGAAGTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATC 515
Db 388 TTGCCGAAACCAGCGAGGGGTGGCGCGATCTCCAAAGGGGGAGGTAGGCCGAAC 447
Oy 516 GATCGGATCTTGCAGAAAGCGTGTGATTCAGAGATTAAAGTGTGCTGTGCAAC 575
Db 448 GGTGCAAGGCGCTGCAGAAAGGGTTGACTTACGGGATTCAACTGCCGGTGTGGAACT 507
Oy 576 TCTACTGCGCTTAACTGGTACTCGACAAACACACTTGCACATATGACTACAAAGCG 635
Db 508 TGTACTGCGCACTCCACCGCTACTCCGACAGCAGACTCAAGTTGACTACCGACTG 567
Oy 636 CAGGCGAGGAAGCGATTGGAAAGCTAATCCTTGTGTGCGCGAGAAAGGTTGTCAAGT 695
Db 568 CTGCCAGGAGCGCATGTGCAAGCTAATCCGTTGTGAGGCAAGCAAGCTCGACAAAGA 627
Oy 696 TTTGATGAGCATCCGTTAAGCTT 718
Db 628 TCTAGGGGGGTTCCCTACGGTT 650
```

Search completed: December 8, 2005, 07:58:35
Job time : 556 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2005, 07:22:03 ; Search time 3730 Seconds
(without alignments)
10084.937 Million cell updates/sec

Title: US-10-716-089-12

Perfect score: 804

Sequence: 1 atccgcggacacgacgcccgc.....agtcgcatcagagctcgcc 804

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	6.1	78.5	755	3	BJ958521 BJ958521
C 2	582.4	72.4	769	3	BJ610693 BJ610693
C 3	582	72.4	733	3	BJ592791 BJ592791
C 4	581	72.3	729	3	BJ592532 BJ592532
C 5	580.4	72.2	747	3	BJ584946 BJ584946
C 6	579.4	72.1	727	3	BJ585055 BJ585055
C 7	570	70.9	721	3	BJ585014 BJ585014
C 8	553	68.8	702	3	BJ606134 BJ606134
C 9	551	68.5	704	3	BJ591761 BJ591761
C 10	490.8	61.0	664	3	BJ65418 BJ65418
C 11	451.4	56.1	642	3	BJ602584 BJ602584
C 12	334.4	42.9	719	3	BJ941820 BJ941820
C 13	334.4	41.6	738	3	BJ605573 BJ605573
C 14	330.4	41.1	736	3	BJ586048 BJ586048
C 15	329.8	41.0	723	3	BJ598461 BJ598461
C 16	307.4	38.2	723	3	BJ611574 BJ611574
C 17	303	37.7	672	3	BJ961154 BJ961154
C 18	292	36.3	670	3	BJ590372 BJ590372
C 19	291	36.2	652	3	BJ952449 BJ952449
C 20	290	36.1	651	3	BJ961357 BJ961357
C 21	287	35.7	589	3	BJ173073 BJ173073
C 22	282	35.1	612	5	BU052287 BU052287

C 23	275.8	34.3	644	3	BJ601969 BJ601969
C 24	274.8	34.2	644	3	BJ597686 BJ597686
C 25	274.8	34.2	644	3	BJ599109 BJ599109
C 26	272.4	33.9	676	3	BJ947687 BJ947687
C 27	267.4	33.3	651	3	BJ602208 BJ602208
C 28	248	30.8	606	3	BJ167125 BJ167125
C 29	231.8	28.8	708	3	BJ590422 BJ590422
C 30	229.4	28.5	726	3	BJ950225 BJ950225
C 31	221.2	27.5	693	3	BJ940325 BJ940325
C 32	199.4	24.8	543	3	BQ827017 BQ827017
C 33	193.6	24.1	609	5	BQ827549 BQ827549
C 34	166.2	20.7	601	3	BJ203788 BJ203788
C 35	166	20.6	571	3	BJ180786 BJ180786
C 36	165	20.5	570	3	BJ196010 BJ196010
C 37	164.8	20.5	543	3	BJ195310 BJ195310
C 38	161.8	20.1	323	3	BJ171535 BJ171535
C 39	161.2	20.0	555	3	BJ192260 BJ192260
C 40	155.4	19.3	554	3	BJ187473 BJ187473
C 41	154.2	19.2	520	3	BJ194455 BJ194455
C 42	150.8	18.8	340	3	BJ163180 BJ163180
C 43	138.8	17.3	515	3	BJ205046 BJ205046
C 44	131.6	16.4	488	3	BJ163628 BJ163628
C 45	129.6	16.1	527	3	BJ205289 BJ205289

ALIGNMENTS

RESULT 1
LOCUS BJ958521/c
DEFINITION BJ958521 pphf full-length cDNA library Physcomitrel:
patens cDNA clone pphf36m05 3', mRNA sequence.
ACCESSION BJ958521
VERSION BJ958521.1 GI:67698288
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
REFERENCE Bukarjota; Viridiplantae; Streptophyta; Embryophyta
Bryopsida; Funariidae; Funariales; Funariaceae; Phy
1 (bases 1 to 755)
AUTHORS Fujita, T., Nishiyama, T., Shin-I, T., Kohara, Y. and H
TITLE Physcomitrella patens EST at a stage of the first a
JOURNAL division of protoplasts
COMMENT unpublished (2005)
Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genres.nig.ac.jp
Protonemata were isolated from the BCDATG medium for ev
Protoplasts were isolated from the protonemata, fur
at 25C under continuous light for 2-3 days. The reg
which were rich in cells at a stage during the fir
cell division, were collected. Total RNA was extrac
constructing a full-length cDNA library. The databa
clones is available at the PHYSCOBASE (<http://moss>
Location/Qualifiers
1..755
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphf36m05"
/issue_type="regenerated protoplasts (ch
/dev_stage="at the first asymmetric cell d
protoplasts"
/clone_lib="pphf full-length cDNA library"
/note="Protonemata were inoculated on BCD,
every ca. 5 days. Protoplasts were isolat
protonemata, further incubated at 25C und

Db 409 CCTCTACTGCGCTTTTACATCGTACTCGAGCAAAACACCTTGACATATGACTACCAAGC 350

Qy 634 CCGAGGGCGAGAAAGCATTGCGAAAGCTAATCTCTTGTGCGCGCGAGAGGTTGTCAA 693

Db 349 CCGAGGGCGAGAAAGCATTGCGAAAGCTAATCTCTTGTGCGCGCGAGAGGTTGTCAA 290

Qy 694 GTTTTGATGAGCATCGCTTAAGCTTTTGTGCGAGAGATTAGGCTTCATATGAGTAA 753

Db 289 GTTTTGATGAGCATCGCTTAAGCTTTTGTGCGAGAGATTAGGCTTCATATGAGTAA 230

Qy 754 CTCTACATCTTTCTTTTATGAGAGAGAGAGTCCATCAAGA 797

Db 229 CTCTACATCTTTCTTTTATGAGAGAGAGAGTCCATCAAGA 186

RESULT 3
 BU592791/c 733 bp mRNA linear EST 22-OCT-2003
 LOCUS BU592791 normalized full length cDNA library, chloronemata, patens
 DEFINITION CDNA clone pphb6121 3', mRNA sequence.
 ACCESSION BU592791 GI:37834779
 VERSION BU592791.1 GI:37834779
 SOURCE EST.
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 733)
 Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M.
 Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
 Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 12808149

JOURNAL Contact: Tadasu Shin-i
 PUBMED Center for Genetic Resource Information
 COMMENT National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tehin@genes.nig.ac.jp

TITLE A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phase vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'-gagagagagagagatccaccccttgagagatgttttttttttttttva-3' was used as a 1st 3' primer, and 5'-gggtctgagtcacgtctgttcagacagcagatgactcgagaccggnnnn-3' as 2nd 5'-hairpin primer, giving the following 5' border sequence, AGGCGAATCGCGAGAGTCCGATTCGCGAGACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on the BCGATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCODbase (<http://meso.nibb.ac.jp/>).
 Location/Qualifiers
 1. 733
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb6121"
 /csize="mixture of chloronemata, caulonemata and malformed buds"
 /clone_lib="normalized full length cDNA library,"

FEATURES
 SOURCE

ORIGIN chloronemata, caulonemata and malformed bu

Query Match 72.4%; Score 582; DB 3; Length 73;
 Best Local Similarity 100.0%; Pred. No. 3e-168;
 Matches 582; Conservative 0; Mismatches 0; Indels 0;

Qy 216 TCGGCGACCAAGCTACCATGAGGCTTGTGCTGAGAGTCTCAACGAGAGAC 275

Db 733 TCGGCGACCAAGCTACCATGAGGCTTGTGCTGAGAGTCTCAACGAGAGAC 674

Qy 276 CGAAGTACGCGCTTTAGCTGAGCAAGCCACTCAGGCTGTCAGGCGAT 335

Db 673 CGAAGTACGCGCTTTAGCTGAGCAAGCCACTCAGGCTGTCAGGCGAT 614

Qy 336 CTGCTGCTGTTTCAGGCCCCCGCTCTCTGATCATGAGACCAAGCTCATAT 395

Db 613 CTGCTGCTGTTTCAGGCCCCCGCTCTCTGATCATGAGACCAAGCTCATAT 554

Qy 336 GAAATGATGTTTGTGCGCATCATCTTCAGCTATCAACAAGACCTGG 455

Db 553 GAAATGATGTTTGTGCGCATCATCTTCAGCTATCAACAAGACCTGG 494

Qy 456 GAGTCCCGCTCAGGCGAGTGAAGTCTCATGCGAGCTCCCTCTAGAC 515

Db 493 GAGTCCCGCTCAGGCGAGTGAAGTCTCATGCGAGCTCCCTCTAGAC 434

Qy 516 GATGCGATCTTTCAGAGAACGCTGTGATGACAGATTTAAAGTGC 575

Db 433 GATGCGATCTTTCAGAGAACGCTGTGATGACAGATTTAAAGTGC 374

Qy 576 TCTACTGCGCTTTTACATGCTGTACTCGAACAAACACACTTGCACATATGA 635

Db 373 TCTACTGCGCTTTTACATGCTGTACTCGAACAAACACACTTGCACATATGA 314

Qy 636 CAGGCGAGAGCGATTCGAGAAAGCTATCCCTTGTGTGCGCGAGAC 695

Db 313 CAGGCGAGAGCGATTCGAGAAAGCTATCCCTTGTGTGCGCGAGAC 254

Qy 636 TTTGATGAGCATCCGTTTAAAGCTTTTCTCCGACGATTTAGGCTTCATAC 755

Db 253 TTTGATGAGCATCCGTTTAAAGCTTTTCTCCGACGATTTAGGCTTCATAC 194

Qy 756 CTACATCTTTCTTTTATCGAGAGCGAGTCCGATCAAGA 797

Db 193 CTACATCTTTCTTTTATCGAGAGCGAGTCCGATCAAGA 152

RESULT 4
 BU592532/c 729 bp mRNA linear
 LOCUS BU592532 normalized full length cDNA library, chloronemata and malformed buds Physcomitrella patens
 DEFINITION CDNA clone pphb6103 3', mRNA sequence.
 ACCESSION BU592532
 VERSION BU592532.1 GI:37834520
 SOURCE EST.
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 729)
 Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Kohara, Y. and Hasebe, M.
 Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
 Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 12808149

JOURNAL Contact: Tadasu Shin-i
 PUBMED Center for Genetic Resource Information
 COMMENT National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- gAgaGAgAgaGATCCACCTTgAgaGATTTTTTTTTTTTTTAA-3' was used as a 1st 3' primer, and 5'- ggttctgagtcATCgctgttccAGACAGcGATgACTgAaACCGNNNN-3' as 2nd 5' hairpin primer, giving the following 5' boarder sequence, AGGCCAATCGCGCGAGCTGCAATTCGCGAGACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13- 14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCOBASE (<http://mos.nibb.ac.jp/>).
Location/Qualifiers
1. 729
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhb6h03"
/tissue_type="mixture of chloronemata, caulonemata and malformed buds"
/clone_lib="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"

Query Match	72.3%;	Score 581;	DB 3;	Length 729;
Best Local Similarity	100.0%;	Pred. No. 6.1e-168;		
Matches 581;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Db	217	CGGCAGCCAGTACCATCGGCTGTGCTGCAAGTCTACCGAGAGACATGCAAGC	276	
Qy	277	GAAGATGACGGTTTATGCTGAGACCAAGCACTGCTGCGGAGATCTGCCACGC	336	
Db	669	GAAGATGACGGTTTATGCTGAGACCAAGCACTGCTGCGGAGATCTGCCACGC	610	
Qy	337	TGCTGCTGTTACGCCCCCGCTCTGTACATGAGACCAAGCTCACATGCGAGTTGAGG	396	
Db	609	TGCTGCTGTTACGCCCCCGCTCTGTACATGAGACCAAGCTCACATGCGAGTTGAGG	550	
Qy	397	AACATGATTTGCGCGCATCAATCTTCCAGCTATCAACAAGCTGTTATCCCGCTGC	456	
Db	549	AACATGATTTGCGCGCATCAATCTTCCAGCTATCAACAAGCTGTTATCCCGCTGC	490	
Qy	457	AGCTGCCCCCTCAGGAGTGAAGTCTCTATGCGAGCTCCCTTAAGCCGAGCCCAATCG	516	
Db	489	AGCTGCCCCCTCAGGAGTGAAGTCTCTATGCGAGCTCCCTTAAGCCGAGCCCAATCG	430	
Qy	517	ATGCGGATCTTGCAAGAGCGTGTGATGAGCAAGATTTAAGTGTGCGCAACT	576	
Db	429	ATGCGGATCTTGCAAGAGCGTGTGATGAGCAAGATTTAAGTGTGCGCAACT	370	
Qy	577	CTACTGCGCTTTACATCGGTACTCGGACAAACACTTGACATATGACTCAAAAGCCGC	636	
Db	369	CTACTGCGCTTTACATCGGTACTCGGACAAACACTTGACATATGACTCAAAAGCCGC	310	
Qy	637	AGGGAGAGAGGATTTGCGAAGCTAATCTTGTGCGGCGCGAAGGTTGCAAGT	696	
Db	309	AGGGAGAGAGGATTTGCGAAGCTAATCTTGTGCGGCGCGAAGGTTGCAAGT	250	
Qy	697	TTGATGAGCATCCGTTAAAGCTTTTCTGCGAGATTTAGGCTTCAATGATTAATCTC	756	

Db	249	TTGATGAGCATCCGTTAAGCTTTTCTGCGAGATTTAGGCTTCAATCA	190
Qy	757	TACATCTTTCTTTTATCGAGAGAGCGATCGCATCAGA	797
Db	189	TACATCTTTCTTTTATCGAGAGAGCGATCGCATCAGA	149

RESULT 5
BU584946/c
LOCUS
DEFINITION
BU584946 normalized full length cDNA library, chloronemata and malformed buds Physcomitrella patens
BU584946
BU584946.1 GI:37826934
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT

Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- gAgaGAgAgaGATCCACCTTgAgaGATTTTTTTTTTTTTTAA-3' was used as a 1st 3' primer, and 5'- ggttctgagtcATCgctgttccAGACAGcGATgACTgAaACCGNN-3' as 2nd 5' hairpin primer, giving the following 5' boarder sequence, AGGCCAATCGCGCGAGCTGCAATTCGCGAGACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13- 14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCOBASE (<http://mos.nibb.ac.jp/>).
Location/Qualifiers
1. 747
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhb26K21"
/tissue_type="mixture of chloronemata, caulonemata and malformed buds"
/clone_lib="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"

Query Match	72.2%;	Score 580.4;	DB 3;	Length 747;
Best Local Similarity	99.7%;	Pred. No. 9.4e-168;		
Matches 581;	Conservative 0;	Mismatches 2;	Indels 0;	

QY 215 TTCGGAGCCAAAGCTACATGAGGATTGTGCTGGAAGTCTACCGAGAGACATGATGCA 274
 Db 747 TTCGGAGCCAAAGTTCATGAGGATTGTGCTGGAAGTCTACCGAGAGACATGATGCA 688
 QY 275 GCGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGCATCTGCGACA 334
 Db 687 GCGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGCATCTGCGACA 628
 QY 335 GCTGCTGCTTTAGAGCCCCCGCTCTCTGATCATGAGACCAAGCTCATCATGCGAGTTGAG 394
 Db 627 GCTGCTGCTTTAGAGCCCCCGCTCTCTGATCATGAGACCAAGCTCATCATGCGAGTTGAG 568
 QY 395 AGAACAATGATTTGTGCGCATCATCTTCAGCTATCAACAAGACCTGGTTACCCCGCT 454
 Db 567 AGAACAATGATTTGTGCGCATCATCTTCAGCTATCAACAAGACCTGGTTACCCCGCT 508
 QY 455 GCAAGTCCCTTCAGGCAAGTGAAGTCTCTATGAGCACTCCCTATGAGCCGAGCCCAAT 514
 Db 507 GCAAGTCCCTTCAGGCAAGTGAAGTCTCTATGAGCACTCCCTATGAGCCGAGCCCAAT 448
 QY 515 CGATGGGATCTTGGCGAGAAAGCTGTGATGACAGATTAAAGTGTGCTGTGCAAC 574
 Db 447 CGATGGGATCTTGGCGAGAAAGCTGTGATGACAGATTAAAGTGTGCTGTGCAAC 388
 QY 575 CTCTACTGGCTTTATCATGCTGATCTGCGAACAAACACTTGCACATATGACTACAAAGCC 634
 Db 387 CTCTACTGGCTTTATCATGCTGATCTGCGAACAAACACTTGCACATATGACTACAAAGCC 328
 QY 635 GCAAGGCGAAGAAAGCAATGCGAAAGCTAACTCTTGTGCGCCGAGAAAGTTGTCAAG 694
 Db 327 GCAAGGCGAAGAAAGCAATGCGAAAGCTAACTCTTGTGCGCCGAGAAAGTTGTCAAG 268
 QY 695 TTTTATGATGACATCCGTAAAGCTTTTCTGCGACGATTTAAGCTTATCATATTGACTAAC 754
 Db 267 TTTTATGATGACATCCGTAAAGCTTTTCTGCGACGATTTAAGCTTATCATATTGACTAAC 208
 QY 755 TCTACATCTTTCTTTATTCGAGAGAGGAGTGGATCAAG 797
 Db 207 TCTACATCTTTCTTTATTCGAGAGAGGAGTGGATCAAG 165

RESULT 6
 BJ585055/c 727 bp mRNA linear EST 22-OCT-2003
 LOCUS BJ585055 normalized full length cDNA library, chloronemata,
 DEFINITION cautionemata and malformed buds Physcomitrella patens subsp. patens
 cDNA clone pphb26n05 3', mRNA sequence.
 ACCESSION BJ585055
 VERSION BJ585055.1 GI:37827043
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 727)
 Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide, H.,
 Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K.,
 Kohara, Y. and Hasebe, M.
 Comparative genomics of Physcomitrella patens gametophytic
 transcriptome and Arabidopsis thaliana: implication for land plant
 evolution
 Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 12808149
 Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS),
 that was in vivo excised from a 1-FLC phage vector (Carninci et al.
 2001). 5' end of the cDNA that was digested with XhoI was ligated

to SalI site of the vector and the 3' end including
 ligated to BamHI site of the
 vector(5'- gagagagagagatccaccctggagagatgtttttttttt
 used as a 1st 3' primer, and
 5'-ggttctcgatgcattcgtctgttcacagacagcgatgactcgaaacccnni
 5'-hairpin primer, giving the following 5' boarder
 AGGCCAATCGCCGACGCTCGAATTCGTCGAAACCG). cDNA inser:
 amplified with conventional T7 and T3 primers. This
 cDNA library was generated according to the method
 Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then
 the BODATG medium for 13-14 days under the contin-
 These clones are available from RIKEN Bio Resource
 (http://www.brc.riken.go.jp/lab/epd/eng/index.html)
 of Physcomitrella EST clones is available at the PH
 (http://moses.nibb.ac.jp).
 Location/Qualifiers
 1..727
 /organism="Physcomitrella patens subsp. pa
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb26n05"
 /tissue_type="mixture of chloronemata, cau
 /clone_lib="normalized full length cDNA li
 chloronemata, cautionemata and malformed bu

ORIGIN
 Query Match 72.1%; Score 579.4; DB 3; Length 727
 Best Local Similarity 99.8%; Pred. No. 1.9e-167;
 Matches 580; Conservative 0; Mismatches 1; Indels 0;

QY 217 CGGACGCCAAGCTACCATGAGGTTGTGCTGGAAGTCTACCGAGAGAC 276
 Db 727 CGGACGCCAAGCTACCATGAGGTTGTGCTGGAAGTCTACCGAGAGAC 668
 QY 277 GAAAGTGAACGGCTTTAAGCTGAGCAAGCCACTCAGGCTGCTCAGGAGAC 336
 Db 667 GAAAGTGAACGGCTTTAAGCTGAGCAAGCCACTCAGGCTGCTCAGGAGAC 608
 QY 337 TGCTGCTGTTAGCCCCCGCTCTGTATCATGAGCAAGCTCACATG 396
 Db 607 TGCTGCTGTTAGCCCCCGCTCTGTATCATGAGCAAGCTCACATG 548
 QY 397 AACCAATGATTGTGCGCATCAATCTTCCAGCTATCAACAAGACTGTT 456
 Db 547 AACCAATGATTGTGCGCATCAATCTTCCAGCTATCAACAAGACTGTT 488
 QY 457 AGCTGCCCTTCAGGCAAGTGAAGTCTCTATGCGAGCTCCCTTAAGACC 516
 Db 487 AGCTGCCCTTCAGGCAAGTGAAGTCTCTATGCGAGCTCCCTTAAGACC 428
 QY 517 AAGCGGATCTGACGAAAGCGTTGATGATGACAGATTAAAGTGGC 576
 Db 427 AATCGGATCTGACGAAAGCGTTGATGATGACAGATTAAAGTGGC 368
 QY 577 CTATGCGCTTTACATCGGTAATCTCGACAAACACACTTGCACATATGAC 636
 Db 367 CTATGCGCTTTACATCGGTAATCTCGACAAACACACTTGCACATATGAC 308
 QY 637 AGGCGAAGAGCGATTTGGAAAGCTAATCTTGTGTCGCGAGAA 696
 Db 307 AGGCGAAGAGCGATTTGGAAAGCTAATCTTGTGTCGCGAGAA 248
 QY 697 TTGATGAGATCCGTTAAGCTTTTCTGCGAGATTTAAGGCTTATAC 756
 Db 247 TTGATGAGATCCGTTAAGCTTTTCTGCGAGATTTAAGGCTTATAC 188
 QY 757 TACATCTTTCTTTATTCGAGAGAGCGAGTGCATCAAGA 797
 Db 187 TACATCTTTCTTTATTCGAGAGAGCGAGTGCATCAAGA 147

RESULT 7
BJS85014/c
LOCUS BJS85014 721 bp mRNA linear EST 22-OCT-2003
DEFINITION BJS85014 normalized full length cDNA library, chloronemata, caulonemata and malformed buds Physcomitrella patens subsp. patens cDNA clone pphb26504 3', mRNA sequence.
ACCESSION BJS85014
VERSION BJS85014.1 GI:37827002
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE 1 (bases 1 to 721)
Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe,M.
Comparative genomes of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
12808149
JOURNAL
PUBMED
COMMENT Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector(5'- gAgaGAGAGAGATCCAAACCCCTGAGAgATTTTATTTTATTTN-3' was used as a 1st 3' primer, and 5'-ggTTCGAGTCATCGCTGTTCAGACAGATGATGACCGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCGCGAGCTGATTCGCGAACC). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCobase (<http://moss.nibb.ac.jp/>).
FEATURES
source location/Qualifiers
1..721
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphb26504"
/ligase_type="mixture of chloronemata, caulonemata and malformed buds"
/clone_1lb="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"
ORIGIN
Query Match 70.9%; Score 570; DB 3; Length 721;
Best Local Similarity 99.7%; Pred. No. 1,5e-164;
Matches 570; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 226 AGCTACCATGGGCTGTGCTCGAAGTGTACCGAGAGACATGACGCAAGATGAC 285
Db 721 ACTACACATGGGGTTGTGCTGTAAGTGNATCCGAGAGACATGACGCAAGATGAC 662
QY 286 GCGTTTATGCTGAGCAAGCATCTAGGCTGCTCAGCGACATCTGCCACAGCTGCTGT 345
Db 661 GCGTTTATGCTGAGCAAGCATCTAGGCTGCTCAGCGACATCTGCCACAGCTGCTGT 602

QY 346 TCAGCCCCCGGCTCTCTGTACATGAGACCAAGCTCATGCGAGGTTGAG; 405
Db 601 TCAGCCCCCGGCTCTCTGTACATGAGACCAAGCTCATGCGAGGTTGAG. 542
QY 406 TGTGCGCATCATCTTCTACAGCTATCAACAAGACCTGTGTTACCCCGCT 465
Db 541 TGTGCGCATCATCTTCTACAGCTATCAACAAGACCTGTGTTACCCCGCT 482
QY 466 TCAGCACTGAAGTCTTATGCGAGCTCCTCTTAGACCCGAGCCCAT 525
Db 481 TCAGCACTGAAGTCTTATGCGAGCTCCTCTTAGACCCGAGCCCAT 422
QY 526 TTGCGAAGAGGCTGTGATTTGACAGATTTTAAGTGTGCTGTGCAAC 585
Db 421 TTGCGAAGAGGCTGTGATTTGACAGATTTTAAGTGTGCTGTGCAAC 362
QY 586 TTATCATCGTACTCGACAAACACACTTGACATATGATACAAAGCC 645
Db 361 TTATCATCGTACTCGACAAACACACTTGACATATGATACAAAGCC 302
QY 646 AGCGATTGCGAAAGTAAATCTCTTGTGTGCGCGAGAAAGTTGTCAAG 705
Db 301 AGCGATTGCGAAAGTAAATCTCTTGTGTGCGCGAGAAAGTTGTCAAG 242
QY 706 ATCCGTTAAGCTTTTCTGCGACGATTTAGGCTTCATCATTTAGTTAAC 765
Db 241 ATCCGTTAAGCTTTTCTGCGACGATTTAGGCTTCATCATTTAGTTAAC 182
QY 766 CTTCTTTATCGAGAGAGCGAGTGGCATCAAGA 797
Db 181 CTTCTTTATCGAGAGAGCGAGTGGCATCAAGA 150
RESULT 8
BJS60134/c
LOCUS BJS60134 702 bp mRNA linear
DEFINITION BJS60134 normalized full length cDNA library, chlor. caulonemata and rhizoid-like protonemata Physcomitrella subsp. patens cDNA clone pphn35f21 3', mRNA sequence
ACCESSION BJS60134
VERSION BJS60134.1 GI:37848126
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Pny: 1 (bases 1 to 702)
Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Kohara,Y. and Hasebe,M.
Comparative genomes of Physcomitrella patens gamet transcriptome and Arabidopsis thaliana: implication evolution
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
12808149
JOURNAL
PUBMED
COMMENT Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector(5'- gAgaGAGAGAGATCCAAACCCCTGAGAgATTTTATTTTATTTN-3' was used as a 1st 3' primer, and 5'-ggTTCGAGTCATCGCTGTTCAGACAGATGATGACCGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCGCGAGCTGATTCGCGAACC). cDNA insert

ant

al.
was

d

amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Proconama[®] were blended by the POLYTRON[®] and then cultivated on the BCDATG medium for 13–14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcometrela EST clones is available at the PHYSCoBase (<http://moses.nibb.ac.jp>).

FEATURES

```

1. .702
/organism="Phycomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_species="patens"
/db_xref="taxon:145481"
/clone="pPhn35f21"
/tissue_type="mixture of chloronemata, callonemata and rhizoid-like protonemata"
/clone_lib="normalized full length cDNA library, chloronemata, callonemata and rhizoid-like protonemata"

```

ORIGIN

Query Match	68.8%	Score 553	DB 3	Length 702
Best Local Similarity	100.0%	Pred. No. 2.8e-159		
Matches 553; Conservative	0	Mismatches 0	Indels 0	Gaps 0

Qy	245	TCGAAGTCTACCCAGAGACAGTCAATGCAAGCAAGATGACGGCTTTAGCTGAGCAAGC	304
Db	702	TCGAAGTCTACCCAGAGACAGTCAATGCAAGCAAGATGACGGCTTTAGCTGAGCAAGC	643
Qy	305	ACTCAGGCTGCTCAGGCGACATCTGCCACAGCTGCTGCTTTACAGCCCCCGCTCTGTGA	364
Db	642	ACTCAGGCTGCTCAGGCGACATCTGCCACAGCTGCTGCTTTACAGCCCCCGCTCTGTGA	583
Qy	365	CATGAGACCAAAGCTTCACATGCGAGTTGAGAGAACATGATTTGTGCCGATCAATCTTCC	424
Db	582	CATGAGACCAAAGCTTCACATGCGAGTTGAGAGAACATGATTTGTGCCGATCAATCTTCC	523
Qy	425	AGCTATACAAAGAAGCTGGTTACCCCGCTGAGAGTGCCTCCAGGCAAGTGAAGTCTCT	484
Db	522	AGCTATACAAAGAAGCTGGTTACCCCGCTGAGAGTGCCTCCAGGCAAGTGAAGTCTCT	463
Qy	485	ATCGACGCTCCCTCTAGACCCGAGCCCAATCGATGCGAGCTTGCAGGAAGCGCTTGTGA	544
Db	462	ATCGACGCTCCCTCTAGACCCGAGCCCAATCGATGCGAGCTTGCAGGAAGCGCTTGTGA	403
Qy	545	TTGACAGAGATTTAAGTGTGCTGTGGCAACTCTAATGCGCTTTTACATCGGTACTCGAGC	604
Db	402	TTGACAGAGATTTAAGTGTGCTGTGGCAACTCTAATGCGCTTTTACATCGGTACTCGAGC	343
Qy	605	AAACACACTTGGCAATATATGACTACAAAGCCGACGGGACGGAAGGATTTGCCAAGACTAAT	664
Db	342	AAACACACTTGGCAATATATGACTACAAAGCCGACGGGACGGAAGGATTTGCCAAGACTAAT	283
Qy	665	CCCTCTTGTGAGGCGGAGAAAGTGTCAAGTTTGTATGATGAGCATCGGTAAAGCTTTTGTGC	724
Db	282	CCCTCTTGTGAGGCGGAGAAAGTGTCAAGTTTGTATGATGAGCATCGGTAAAGCTTTTGTGC	223
Qy	725	CGAGATTTAGGCTTCATACATTGAGTAACCTCAATCTTTCTTTTATCGAGAGACGG	784
Db	222	CGAGATTTAGGCTTCATACATTGAGTAACCTCAATCTTTCTTTTATCGAGAGACGG	163
Qy	785	AGTCGATCAAGA	797
Db	162	AGTCGATCAAGA	150

RESULT 9

LOCUS	704 bp	mRNA	linear	EST 22-OCT-2003
B5591761/c	B5591761	B5591761	normalized full length cDNA library, chloromemata, cDNAemata and malformed buds Physcomitrella patens subsp. patens	
DEFINITION	cDNA clone pphb4121 3', mRNA sequence.			

ACCESSION	BU591761
VERSION	BU591761.1
	GI:37833749

KEYWORDS

SOURCE ORGANISM	physcomitrella patens subsp. patens physcomitrella patens subsp. patens

REFERENCE
AUTHORS
1 (bases 1 to 704)
Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishida

TITLE

transcriptome and *Arabidopsis thaliana*: implications for evolution

COMMENT

Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tsuhni@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript⁺ that was in vivo excised from a λ-FLC phage vector (2001). 5' end of the cDNA that was digested with XbaI to SalI site of the vector and the 3' end including ligated to BamHI site of the vector (5'- ggaagagagagatcagaatccctggaagatttttttttttt used as a 1st 3' primer, and 5'-ggatttcgaagcatcgctgttccacagacagcagtatgcgaaccggnnaggccaattgcgccagatcgcaatttcgcgaaccg). cDNA insert amplified with conventional T7 and T3 primers. This cDNA library was generated according to the method Nishiyama et al. (2003).

Pichonemata were blended by the POLYTRON[®], and then the BCDATG medium for 13–14 days under the continuous These clones are available from RIKEN Bio Resource (<http://www.brc.riken.go.jp/lab/epd/eng/index.html>) of Physcomitrella EST clones is available at the Phytozoo (<http://mooss.nibb.ac.jp/>).

FEATURES

```

1. .704
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pghb44121"
/tissue_type="mixture of chloronemata, caulonemata buds"
malformed buds"
/clone_libs="normalized full length cDNA library of chloronemata and malformed buds"
/chloronemata, caulonemata and malformed buds"

```

ORIGIN

Query Match	68.5%;	Score 551;	DB 3;	Length 70;
Best Local Similarity	99.8%;	Pred. No. 1.2e-158;		
Matches 551; Conservative	0;	Mismatches 1;	Indels	

QY 246 CGAAGTGTCTACCGAAGACAGTCATGCAACCGAAGATGACGGCTTTAG
Db 704 CGAAGTGTCTACCGAAGACAGTCATGCAACCGAAGATGACGGCTTTAG
QY 306 CTCAAGGTCTCAGGCGCATCTGCGACAGCTGCTGCTTCAGCCCC
Db 644 CTCAGGCTTCTCAGGCGCATCTGCGACAGCTGCTGCTTCAGCCCC
QY 366 ATGAAACCAAGTCATGCGAAGTTGAGAAACAATATTGTGCGCA
Db 584 ATGAGACCAAGTCATGCGAAGTTGAGAAACAATATTGTGCGCA
QY 426 GGTATCAACMAAGCTGTGTTACCCCGGTGCAAGCTGCCCTCAGGCACT

Db 524 GGTATCAACAGACCTGGTTACCCCGCTGACAGTCCCTCCAGGCAAGTGAATCCTCTA 465
QY 466 TCGAGCTCCCTCTGAGACCCGAGCCCAATGCAATGGGAATCTTGAGAGAGCGTTGTGAT 545
Db 464 TCGAGCTCCCTCTGAGACCCGAGCCCAATGCAATGGGAATCTTGAGAGAGCGTTGTGAT 405
QY 546 TGACAGATTTTAAAGTGTGCTGTGCAACCTTACTGCGCTTTTAACTCGGTACTCGACA 605
Db 404 TGACAGATTTTAAAGTGTGCTGTGCAACCTTACTGCGCTTTTAACTCGGTACTCGACA 345
QY 606 AACACACTTGACATATGACTACAAAGCCGACGGGCGAGAGCGAATGGGAAAGCTAATC 665
Db 344 AACACACTTGACATATGACTACAAAGCCGACGGGCGAGAGCGAATGGGAAAGCTAATC 285
QY 666 CTCTGTGTGTCGCGAGAAAGTTGTCAAGTTTGAATGAGACCTGTAAGCTTTCTGTGCC 725
Db 284 CTCTGTGTGTCGCGAGAAAGTTGTCAAGTTTGAATGAGACCTGTAAGCTTTCTGTGCC 225
QY 726 GACGATTTAGGCTTCAATGAGTAAGTCTTACATCTTTCTTTTATCGAGAGAGCGA 785
Db 224 GACGATTTAGGCTTCAATGAGTAAGTCTTACATCTTTCTTTTATCGAGAGAGCGA 165
QY 786 GTGCATCAAGA 797
Db 164 GTGCATCAAGA 153

RESULT 10
BU165418/c 664 bp mRNA linear EST 16-OCT-2003
DEFINITION BU165418 full length cDNA library, chloronemata and young gametophores Physcomitrella patens subsp. patens cDNA clone pph4c11 3', mRNA sequence.

ACCESSION BU165418 GI:18333403
VERSION BU165418.1
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
REFERENCE Bryophyta; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 664)
AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe,M.
Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution

TITLE Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
JOURNAL 12808149
PUBMED Contact: Tadao Shin-I
COMMENT Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a l-PUC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was vector(5'- gagagaagagatccacacctgagagattttttttttttttttn-3' was used as a 1st 3' primer, and 5'-ggttcctgagtcacgtcgttcctgagacgcatgacgaaacggnnnn-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCGCAATCGCGAGCTCGAATTCGCGAACC). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database

of Physcomitrella EST clones is available at the PF (<http://mos.nibb.ac.jp>).

FEATURES
source Location/Qualifiers
1. 664

/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pph4c11"
/issue_type="mixture of chloronemata and young gametophores with 2 to 5 leaves"
/clone_lib="full length cDNA library, chloronemata and young gametophores"

ORIGIN

Query Match 61.0%; Score 490.8; DB 3; Length 67
Best Local Similarity 99.6%; Pred. No. 4.8e-140;
Matches 492; Conservative 0; Mismatches 2; Indels 0;

QY 304 CACTCAGGCTGCTCAGGAGCATCTGCCAGACGCTGCTGTTCAGCCC 363
Db 664 CACTCAGGCTGCTCAGGAGCATCTGCCAGACGCTGCTGTTCAGCCC 605
QY 364 ACATGAGACCAAGCTCAGATGAGAGTGAAGAACTATGTGCGC 423
Db 604 ACATGAGACCAAGCTCAGATGAGAGTGAAGAACTATGTGCGC 545
QY 424 CAGCTATCAACAAAGCTGTGTTACCCCGCTGACGCTCCCTCAGGCA 483
Db 544 CAGCTATCAACAAAGCTGTGTTACCCCGCTGACGCTCCCTCAGGCA 485
QY 484 TATGCACTCCCTCTAGACCCGAGCCCAATGATGAGGATCTTGCAGG 543
Db 484 TATGCACTCCCTCTAGACCCGAGCCCAATGATGAGGATCTTGCAGG 425
QY 544 ATTGACAGATTTAAGTGTGCTGTGAGCACTTACTGCGCTTTACAT 603
Db 424 ATTGACAGATTTAAGTGTGCTGTGAGCACTTACTGCGCTTTACAT 365
QY 604 CAAACACACTTGACATATGACTACAAAGCCGACGGGCGAGAGAGGAT 663
Db 364 CAAACACACTTGACATATGACTACAAAGCCGACGGGCGAGAGAGGAT 305
QY 664 TCCTCTGTGTCGCGCGAGAAAGTTGTCAAGTTTGTAGAGCATCGTT 723
Db 304 TCCTCTGTGTCGCGCGAGAAAGTTGTCAAGTTTGTAGAGCATCGTT 245
QY 724 CCGACGATTTAGGCTTCAATGAGTAAGTCTTACATCTTTCTTTT 783
Db 244 CCGACGATTTAGGCTTCAATGAGTAAGTCTTACATCTTTCTTTT 185
QY 784 GAGTCGATCAAGA 797
Db 184 GAGTCGATCAAGA 171

RESULT 11

BU602584/c

LOCUS BU602584 642 bp mRNA linear
DEFINITION BU602584 normalized full length cDNA library, chloronemata and rhizoid-like protonemata Physcomitrella subsp. patens cDNA clone pph4a13 3', mRNA sequence

ACCESSION BU602584 GI:37844576
VERSION BU602584.1
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
REFERENCE Bryophyta; Viridiplantae; Streptophyta; Embryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 642)
AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y., Kohara,Y. and Hasebe,M.

K.,

TITLE

Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
PUBMED
12808149

COMMENT
Contact: Tadaeu Shin-1

Center For Genetic Resource Information

National Institute of Genetics

111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-PLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'-gagagagagagagatccacacctgagagattttttttttttttt-3' was used as a 1st 3' primer, and 5'-gggtctcgagtcagtcgttccagacagcgatgacgagacccgmn-3' as 2nd 5'-hairpin primer, giving the following 5' leader sequence, AGGCCAATCCGCGGAGCTCGAATTCGTGGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of *Physcomitrella* EST clones is available at the PHYSCODbase (<http://moses.nibb.ac.jp>).

FEATURES
source

1..642
/location/Qualifiers
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhn8a13"
/tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"
/clone_1lb="normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN

Query Match 56.1%; Score 451.4; DB 3; Length 642;
Best Local Similarity 99.3%; Pred. No. 7,5e-128;

Matches 452; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

343 TGTTCAGCCCCCGCTCTGTACATGAGACCAAGCTCAGTGGATTGAGAGAACAT 402
642 TGTTCAGCCCCCGCTCTGTACATGAGACCAAGCTCAGTGGATTGAGAGAACAT 583
403 GATTGTGCGCATCAATCTTCAGCTATCAACAGACCTGTTACCCCGCTGACAGTGC 462
582 GATTGTGCGCATCAATCTTCAGCTATCAACAGACCTGTTACCCCGCTGACAGTGC 523
463 CCCCTCAGCAGTGAAGTCTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATGATGCGG 522
522 CCNTCAGCAGTGAAGTCTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATGATGCGG 463
523 ATCTTCAGGAAGCGTGTGATTGATGACGAGATTAAAGTGTGCTGTGCAACCTTACTG 582
462 ATCTTCAGGAAGCGTGTGATTGATGACGAGATTAAAGTGTGCTGTGCAACCTTACTG 403
583 CGCTTTACATCGGTACTCGGACAAACACATTCGACATATGCTACAAAGCGGAGGCA 642
402 CGCTTTACATCGGTACTCGGACAAACACATTCGACATATGCTACAAAGCGGAGGCA 343
643 GGAAGCGATTGCGGAAGTAATCTCTGTTGCTGCGCGGAGAAAGTTGTCAAGTTTGTATG 702
342 GGAAGCGATTGCGGAAGTAATCTCTGTTGCTGCGCGGAGAAAGTTGTCAAGTTTGTATG 283
703 AGCATCCGTTAAGTTTCTGCGGAGATTTAGGCTTCAATGAGTAACCTACATC 762

DB 282 AGCATCCGTTAAGTTTCTGCGGAGATTTAGGCTTCAATGAGTAACCTACATC 223

QY 763 TTTCTTCTTTATCGAGAGCGAGTGCATCAAGA 797

DB 222 TTTCTTCTTTATCGAGAGCGAGTGCATCAAGA 188

RESULT 12

BU941820

LOCUS

DEFINITION

BU941820 pphf full-length cDNA library Physcomitrella

ACCESSION

BU941820

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Tadaeu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
Protonemata were inoculated on BCDATG medium for ex
protoplasts were isolated from the protonemata, fu
at 25C under continuous light for 2-3 days. The reg
which were rich in cells at a stage during the first
cell division, were collected. Total RNA was extrac
constructing a full-length cDNA library. The databa
clones is available at the PHYSCODbase (<http://moses>).
Location/Qualifiers
1..719
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhf16106"
/tissue_type="regenerated protoplasts (ch)
/dev stage="at the first asymmetric cell d
protoplasts"
/clone_1lb="pphf full-length cDNA library"
/note="Protonemata were inoculated on BCDATG
every ca. 5 days. Protoplasts were isolate
protonemata, further incubated at 25C unde
light for 2-3 days. The regenerated cells,
in cells at a stage during the first asym
division, were collected. Total RNA was ex
constructing a full-length cDNA library."

FEATURES
source

1..719
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhf16106"
/tissue_type="regenerated protoplasts (ch)
/dev stage="at the first asymmetric cell d
protoplasts"
/clone_1lb="pphf full-length cDNA library"
/note="Protonemata were inoculated on BCDATG
every ca. 5 days. Protoplasts were isolate
protonemata, further incubated at 25C unde
light for 2-3 days. The regenerated cells,
in cells at a stage during the first asym
division, were collected. Total RNA was ex
constructing a full-length cDNA library."

ORIGIN

Query Match 42.9%; Score 344.8; DB 3; Length 719;
Best Local Similarity 72.7%; Pred. No. 7.6e-95;

Matches 445; Conservative 0; Mismatches 167; Indels 0;

QY 162 GCAGCTCGAGGCCCTGAGGATCCAGTTATGTCGAAGAACCTTGGCGG 221
DB 1 CAACCTCCAGACCCCGAGGGGCTGTGATGTGTAAAGACGTGTGCGG 60
QY 222 GCCAAGCTACATGGGGTTGTGCTCGAAGTGTCTACCGAGAGACATCAT 281
DB 61 GCCAAGCTACATGGGGTTGTGCTCGAAGTGTCTACCGAGAGACATCAT 120
QY 282 TGAGGGGTTTGTGAGCAAGCACTGAGGCTGCTGAGGCGACATCTGC 341

Db 121 TGACCGCCGTGGCCGAGCAGGCACTCAAGCCGCTCAGGTGCTGCATCCGCTGATCCA 180

Qy 342 CTGTTCAAGCCCCCGCTCTGTACATGAGACCACTCACTGCAAGGTGAGAAACAA 401

Db 181 GTGCTCACCCCTTCAAGTTCTTAAGAGAGAGACAGAGCTCATTTGAGGCCGACAGCATGT 240

Qy 402 TGATGTGCGCGCATCATCTTCCAGCTATCAACAGACCTGGTTACCCCGCTGACGCTG 461

Db 241 TGATCAACCCGCTCAATCTTCCAGCCACATCACTGAGGTTCTCCGTTACAGTAG 300

Qy 462 CCCCTCAGGCAAGTAACTCTCTATCGACGCTCCCTTAGAACCCGACCAATGATSCG 521

Db 301 CCCCCAGGTTGTAGTTGGCCCCGTTGCCACACCTCCAGACCGGACCAACCGGTCG 360

Qy 522 GATCTTGCAGAGAGCGTGTGATGATGACAGATTTAAGTGTGCTGTGCAACCTCTACT 581

Db 361 GATCTTGCAGAGAGCGCGTGTGATGACAGATTTCAAGTGTGCTGTGCAACCTCTTTT 420

Qy 582 GCGCTTTCATGCTGACTCGACAAACACACTTGACATATGACTACAAAGCCGACGGC 641

Db 421 GTGCTTACATGCTGATTTCTGCAACACATATGTTGACATTAATTAACAGGCTGCTGAC 480

Qy 642 AGGAAGCAATTCGAAAGTAATCTCTTGTGCTGCGGAGAGGTTGCAAGTTTGAT 701

Db 481 AGGAGGCTATTCGAAAGCTATCTTTAGTTGTGCGGAGAGGTTGTTAGTTCTGAA 540

Qy 702 GAGCATCCGTTAAGCTTTTCTGCCAGCATTTAGGCTTCATACATTAAGTAACTTAAT 761

Db 541 GAGGAGCATGCTGACCTTTTGTATGTGACGATGAGCTTCAGGATGAATAGTTCTTAAC 600

Qy 762 CTTTCTCTTTA 773

Db 601 TTTCTTTATTGA 612

RESULT 13
 BJ605573/c 738 bp mRNA linear EST 22-OCT-2003
 LOCUS
 DEFINITION BJ605573 normalized full length cDNA library, chlorememata, caulonemata and rhizoid-like protonemata Physcomitrella patens subsp. patens cDNA clone pphn34m03 3', mRNA sequence.

ACCESSION BJ605573.1 GI:37847565
 VERSION
 KEYWORDS EST
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 738)
 AUTHORS Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M.
 REFERENCE Comparative genomes of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
 TITLE Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 JOURNAL 12808149
 PUBMED
 COMMENT Contact: Tadao Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was ended with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'-GAGAGAGAGAGATCCACCTGAGAGATTTTTTTTTTTTNN-3' was used as a 1st 3' primer, and 5'-GGTTCGAGTCATGCTTCAGACAGCGATGATGAGAACGNNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence,

AGGCCAAATGCGCGGAGCTGATTCGAGAGACCG). cDNA insert amplified with conventional T7 and T3 primers. This cDNA library was generated according to the method Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then the BODATS medium for 13-14 days under the continn These clones are available from RIKEN BGI Resource (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>) of Physcomitrella EST clones is available at the PH (<http://mos.nibb.ac.jp>).
 Location/Qualifiers

FEATURES
 source
 1..738
 /organism="Physcomitrella patens subsp. pa"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphn34m03"
 /tissue_type="mixture of chloronemata, cau rhizoid-like protonemata"
 /clone_lib="normalized full length cDNA li" chloronemata, caulonemata and rhizoid-like

ORIGIN
 Query Match 41.6%; Score 334.4; DB 3; Length 7
 Best Local Similarity 72.3%; Pred. No. 1.3e-91;
 Matches 434; Conservative 0; Mismatches 166; Indels 0;

Qy 174 CCCCTGAGGTCCTCACTTATGTCAGAAACCTTTGGCGTCTTCTGGGCGAG 233

Db 738 CCCGAGGGGGCCCTGTGATGTGTAAAGCGTGTGGGTTCTTTGGCAG 679

Qy 234 TGGGGTTGTCTCGAAGTGTACCGAGAGACATGATCAACAGAGAT 293

Db 678 TGGGTCTCTCTCGAAGTGTACCGAGAGACATGATCAACAGAGAT 619

Qy 294 CTGAGCAAGCCACTGAGGCTCTCAGGCGACATCTGCCAGCTGCTGC 353

Db 618 CCGAGCAGGCCACTCAACCGCTCAGGTGCTGCATCCGCTGCATCCAG 559

Qy 354 CCGCTCTGTATGAGACCAAGCTCAATCGAGGTTGAGAGAACAT 413

Db 558 CAGTCTTATGAGAAAGAGCAAGAGCTAATTTGAGGCGGACAGATGT 499

Qy 414 ATCAATCTTCCAGCTATCAACAGACCTGTATACCCCGCTGCAGCTGC 473

Db 498 CTCAATCTTCCAGCCACATCCAGTGTAGGTTGCTCCGTTACATAGC 439

Qy 474 TGAAGTCTCTATGCGACGCTCCCTTAGAACCCGAGCCCAATGATGCG 533

Db 438 TAGTTGCCCGGTTSCACACCCCTCAGACCGGACCAACCGGTGCGG 379

Qy 534 AGCGTTGATGATGACAGATTTAAGTGTGCTGTGCGCAACCTCTACG 593

Db 378 AGCGGTGGGATGACAGATTTTCAGTGTGTTGGGACCTTTTTTGG 319

Qy 594 GGTACTCGAACAACACACTTGACATATGACTACAAAGCCGACAGGCA 653

Db 318 GGTATTTGACAAACATATGTTGCACTTATGATTAAGGCTGCTGACA 259

Qy 654 CGAAAGCTATTCCTTGTGCGGCGGAGAGGTTGCAAGTTTGTATG 713

Db 258 CAAAGGCTATTCCTTGTGTTGGCGGAGAGGTTGTTAAGTTGTAAG 199

Qy 714 AGCTTTCTGCGAGAGATTTAGGCTTCAATCAATTAAGTAACTCTACATC 773

Db 198 GACCTTTTGTATGTGACGATGAGCTTCAGGATGAGAAATAGTTCTTAAC 139

RESULT 14
 BJ586048/c 736 bp mRNA linear
 LOCUS
 DEFINITION BJ586048 normalized full length cDNA library, chlor: caulonemata and malformed buds Physcomitrella patens

CDNA clone pphb29418 3', mRNA sequence.

ACCESSION BJ586048
 VERSION BJ586048.1 GI:37828036
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Bacteria; Viridiplantae; Streptophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 REFERENCE 1 (bases 1 to 736)
 AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinzaki,K., Kohara,Y. and Hasebe,M.
 TITLE Comparative genomes of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 PUBMED 12808149
 COMMENT Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tschini@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-PhiC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- GAGAGAGAGAGATCCAAACCTGAGAGATTTTCTTTTTCVN-3' was used as a 1st 3' primer, and 5'-ggttctgagtcgagtcgcttccgacagcgatgactgagaaacggnnn-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATGCGCCGACTCGAATTCGTCGAAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCObase (<http://moses.nibb.ac.jp>).
 FEATURES
 source location/Qualifiers
 1..736
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb29418"
 /ciseue_type="mixture of chloronemata, caulonemata and malformed buds"
 /clone_1lb="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"

ORIGIN

Query Match 41.1%; Score 330.4; DB 3; Length 736;
 Best Local Similarity 72.6%; Pred. No. 2.2e-90;
 Matches 427; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

186 CAGTATATGCAAGAACTTGGGGCTTCTGGGACCAACTACATGGGGTTGTGT 245
 DB CTGTATATGTAAGAACTGTGGCTTTCTTGGGACCAACTATGGGTTCTGTCT 677
 246 CGAAGTGTACCGAGACAGTCAATGCAAGCAAGATGACGCTTTAGCTGAGCAAGCA 305
 DB CGAAGTGTACCGAGACAGTCAATGCAAGCAAGATGACGCTTTAGCTGAGCAAGCA 617
 306 CTCAGGCTGCTAGGCGACATTTGCCACAGCTGCTGCTTTAGGCCCCCGCTCTTAC 365
 DB CTCAGGCTGCTAGGCGACATTTGCCACAGCTGCTGCTTTAGGCCCCCGCTCTTAC 557
 366 ATGAGACCAAGTCAATGCGAGGTTGAGAGAAATGATTTGTGCGCATCAATCTTCCA 425

DB 556 AAGAGACAGAGCTTCATTTGAGCCGACAGCATGTTGATTAACCGCC
 426 GCTATCAACMAAGCTGTGTTAACCCCGCTGACGCTGCCCTCAGCAG?
 DB 496 GGCACATCAAGTTAGAGTTGTCTCCGGTTACAGTAGTACCCCAAGTTG?
 486 TGGCAGCTCCCTCTTACGCCGAGCCCAATGATGCGGATCTTGCAGGA,
 DB 436 TTGCCACACCCCTCCAGACCGGACCAACCCGCTGGGATCTTGCAGGA,
 546 TCAGAGATTTAAGTGTGCTGGCAACCTCTACTGCGCTTTACATC;
 DB 376 TCACAGATTTCAATGTCGTTGTTGGCACCTTTTGTGCTTTACATC;
 606 AACACATTTGACATATGACTACAAAGCCGACGAGGACGAGAACGATTC;
 DB 316 AACATAGTTGACATTAATGATTAACAGGCTGCTGACAGAGGCTATTGC
 666 CTCTTGTGCTGCGCGAGAAAGTTTCAAGTTTGAATGACATCCGTTA,
 DB 256 CTTTAGTTGTGCGCGAGAAAGTTTGAATGTTGAAAGTTCGAAGGGACTCTC,
 726 GACGATTTAGGCTCATATGATGATGATGATCTACATCTTTCTTTA
 DB 136 GTGACGTRAGGCTTCAGCATGAAATGTTCTTAACCTTTCTTAATGA

RESULT 15
 BJ598461/c
 LOCUS
 DEFINITION
 BJ598461 normalized full length cDNA library, chloronemata and rhizoid-like protonemata Physcomitrella patens subsp. patens
 ACCESSION BJ598461
 VERSION BJ598461
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Bacteria; Viridiplantae; Streptophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 REFERENCE 1 (bases 1 to 736)
 AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Kohara,Y. and Hasebe,M.
 TITLE Comparative genomes of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for evolution
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012
 PUBMED 12808149
 COMMENT Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tschini@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-PhiC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- GAGAGAGAGAGATCCAAACCTGAGAGATTTTCTTTTTCVN-3' was used as a 1st 3' primer, and 5'-ggttctgagtcgagtcgcttccgacagcgatgactgagaaacggnnn-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATGCGCCGACTCGAATTCGTCGAAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource

(<http://www.brc.riken.go.jp/lab/spd/Eng/index.html>). The database of *Physcomitrella* EST clones is available at the PHYCObase (<http://mos.nid.ac.jp/>).

FEATURES

source

Location/Qualifiers

1. . 723

/organism="Physcomitrella patens subsp. patens"

/mol_type="mRNA"

/sub_species="Patens"

/db_xref="taxon:145481"

/clone="pbin23j11"

/tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"

/clone_lib="normalized full length cDNA library,"

chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN

Query Match 41.0%; Score 329.8; DB 3; Length 723;
Best Local Similarity 72.5%; Pred. No. 3.4e-90;
Matches 427; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

```
OY 185 CCAGTTATGTGCAAGAACCTTTGGCGCTTCTTCGCGACCAAGCTACCATGGGTTGTGC 244
    |||||
DB 723 CCTGTGATGTGTAAAGACGTGCGGTTCTTTGGCAGCCAGCCATATGGGTCCTGC 664

OY 245 TCGAAGTCTACCGAGAGCAAGTCATGCAAGCGAAGATGACGGCTTTAGCTGAGCAAGCC 304
    |||||
DB 663 TCGAAGTCTTACAGGAGAACGTCATGCAAGCGAAGATGACGGCGGTGCGAGCAGGCC 604

OY 305 ACTCAGGCTGCTCAGGCGACATCTGCCAAGCTGCTGTTCAGCCCCCGCTCTGTGA 364
    |||||
DB 603 ACTCAAGCCGCTCAGGTGCTGCATCCGTCGATCCAGGCTCAGCTCCAGTTCTATG 544

OY 365 CATGAGCAACAGCTCAATGCGAGGTGAGAGAAACAATGATGTGCGCATCATCTCC 424
    |||||
DB 543 GAAGAGCAACAGAGCTCAATTGAGGCCGACAGCATGTTGATACAAACCGCTCAATCTTCC 484

OY 425 AGCTATCAACAAGACCTGGTTACCCCGCTGCAGCTGCCCTCAGGAGTAAATCTCT 484
    |||||
DB 483 AGCCACCATCCAGTTGAGGTGCTCCGGTTACAGTAGCCCCCAGGTTGATGCCCCC 424

OY 485 ATCGAGCTCCCTCTAGAACCCGAGCCCAATCGATCGGATCTTGCAGGAAGCGTGTGA 544
    |||||
DB 423 GTTGCACACCCCTCCAGACCGGCAACAAACCGGTGCGGATCTTGCAGGAAGCGGTGGA 364

OY 545 TTGACAGATTTAAATGTCGCTGTGGCAACCTCTACTGCGCTTTACATCGGTACTCGAC 604
    |||||
DB 363 TTGACAGATTTCAAGTGTGCTGTGGGACCTTTTGTGCTTTACATCGGTATTCTGAC 304

OY 605 AAACACACTTGCACATATGACTACAAAGCGCAGAGGAGAGAGATTCGAAAGCTAAT 664
    |||||
DB 303 AAACATAGTGTGCACTTATGATTTACAAAGGTGCTGGAACAGAGGCTATTGCCAAAGGCTAAT 244

OY 665 CCTCTTGTGTCGTCGAGAAAGGTTGCAAGTTTGAATGAGCATCCGTTAAGCTTTTCTGC 724
    |||||
DB 243 CCTTTAGTTGTGCGCAGAAAGGTTGTTAAGTTCTGAAGAGGAGCTGCGACCTTTTGA 184

OY 725 CGACGATTTAAGGCTTCATACATTTGAGTAACTCTACATCTTTCTTTA 773
    |||||
DB 183 TGTGACGTAAGGCTTCAGGCAATGAAATAGTCTTAACTTTCTTTATGTA 135
```

Search completed: December 8, 2005, 10:12:06
Job time : 3734 secs

QY 301 AGCACTGAGGCTGCTCAGGCGACATCTGCA CAGTGTGCTGTTCAAGCCCCCGCTCC 360
|
|
|
Db 301 AGCACTGAGGCTGCTCAGGCGACATCTGCA CAGTGTGCTGTTCAAGCCCCCGCTCC 360
|
|
|
QY 361 TGTACATGAGACCAAGCTCAGATGCGAGGTTGAGAGAACATGATTGTGCGGCATCAATC 420
|
|
|
|
Db 361 TGTACATGAGACCAAGCTCAGATGCGAGGTTGAGAGAACATGATTGTGCGGCATCAATC 420
|
|
|
QY 421 TTCCAGCTATCAACAGACCTGTTACCCCGCTGAGTGCCTCCAGGCGTGAATGC 480
|
|
|
|
Db 421 TTCCAGCTATCAACAGACCTGTTACCCCGCTGAGTGCCTCCAGGCGTGAATGC 480
|
|
|
QY 481 CTCTATGCGAGCTCCCTTAGACCCGAGCCCAATGATGCGGATCTTGGAGAGAGGTGT 540
|
|
|
|
Db 481 CTCTATGCGAGCTCCCTTAGACCCGAGCCCAATGATGCGGATCTTGGAGAGAGGTGT 540
|
|
|
QY 541 TGGATTGACAGGATTTAAGTGTGCTGTGCGCAACTCTA CTGCGCTTTACATCGGTACTC 600
|
|
|
|
Db 541 TGGATTGACAGGATTTAAGTGTGCTGTGCGCAACTCTA CTGCGCTTTACATCGGTACTC 600
|
|
|
QY 601 GGACAAACACACTTTGCACTATGACTATCAAAAGCCGAGGGCGAGAAAGCATTCGAAAGC 660
|
|
|
|
Db 601 GGACAAACACACTTTGCACTATGACTATCAAAAGCCGAGGGCGAGAAAGCATTCGAAAGC 660
|
|
|
QY 661 TAATCCTCTTGTGTCGTCGCGAGAGAGGTGTCAAGTTTGTATGAGCATCCGTTAGCTTTT 720
|
|
|
|
Db 661 TAATCCTCTTGTGTCGTCGCGAGAGAGGTGTCAAGTTTGTATGAGCATCCGTTAGCTTTT 720
|
|
|
QY 721 CTGCGGACGATTTAAGCTTCATACATTGAGTAAGTCTTACATCTTTCTTTTATCGAAG 780
|
|
|
|
Db 721 CTGCGGACGATTTAAGCTTCATACATTGAGTAAGTCTTACATCTTTCTTTTATCGAAG 780
|
|
|
QY 781 AGCGAGTGGCATCAAGAGCTCGCC 804
|
|
|
|
Db 781 AGCGAGTGGCATCAAGAGCTCGCC 804
|
|
|

RESULT 2
US-09-828-303-4

/ Sequence 4, Application US/09828303
/ Patent No. 6677504
/ GENERAL INFORMATION:
/ APPLICANT: COSTA E SILVA, OSWALDO DA
/ APPLICANT: BOHNER, HANS J.
/ APPLICANT: VAN THIELEN, NOCHA
/ APPLICANT: CHEN, ROUYING
/ TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
/ TITLE OF INVENTION: METHODS OF USE IN PLANTS
/ FILE REFERENCE: 16313-0030
/ CURRENT APPLICATION NUMBER: US/09/828,303
/ CURRENT FILING DATE: 2001-08-20
/ PRIOR APPLICATION NUMBER: 60/196,001
/ PRIOR FILING DATE: 2000-04-07
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 1531
/ TYPE: DNA
/ ORGANISM: Physcomitrella patens
/ US-09-828-303-4

Query Match 96.6%; Score 776.4; DB 3; Length 1531;
Best Local Similarity 99.7%; Pred. No. 2.1e-244;

Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 8 GCACCAAGTCCCTTAGTGTGTGTCATTAAGTGTGTTGCAAGTCAAGCTTGAGC 67
|
|
|
Db 529 GCACCAAGTCCCTTAGTGTGTGTCATTAAGTGTGTTGCAAGTCAAGCTTGAGC 588
|
|
|
QY 68 GAGATTTCAGAGATTTCATACAGCTTGTGATTGAGAAAGATACACCTTATTAGTCTG 127
|
|
|
|
Db 589 GAGATTTCAGAGATTTCATACAGCTTGTGATTGAGAAAGATACACCTTATTAGTCTG 648
|
|
|

QY 128 TTTAAGATGACCAAGAGCGTGTGTTCTCAGGAGACGACCTCGAGAGCCC 187
|
|
|
Db 649 TTTAAGATGACCAAGAGCGTGTGTTCTCAGGAGACGACCTCGAGAGCCC 708
|
|
|
QY 188 GTTATGTCAAGAACCTTTGCGGCTTCTTGGCGAGCCAAAGCTTACCATGAC 247
|
|
|
|
Db 709 GTTATGTCAAGAACCTTTGCGGCTTCTTGGCGAGCCAAAGCTTACCATGAC 768
|
|
|
QY 248 AAGTCTACCGAGAGACGTCATGCAAGCGAAGATGAGCGCTTTAGCTGT 307
|
|
|
|
Db 769 AAGTCTACCGAGAGACGTCATGCAAGCGAAGATGAGCGCTTTAGCTGT 827
|
|
|
QY 308 CAGGCTGCTCAGGCGACATCTGCAACAGCTGCTGTTCAAGCCCCCG 367
|
|
|
|
Db 828 CAGGCTGCTCAGGCGACATCTGCAACAGCTGCTGTTCAAGCCCCCG 887
|
|
|
QY 368 GAGACCAAGCTCACTGCGAGGTTGAGAGAACATGATTGTGCGCATC 427
|
|
|
|
Db 888 GAGACCAAGCTCACTGCGAGGTTGAGAGAACATGATTGTGCGCATC 947
|
|
|
QY 428 TATCAACAAAGCTGTGTTACCCCGCTGACGTGCTCCCTCAGGCGATG 487
|
|
|
|
Db 948 TATCAACAAAGCTGTGTTACCCCGCTGACGTGCTCCCTCAGGCGATG 1007
|
|
|
QY 488 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTCAAGAGAC 547
|
|
|
|
Db 1008 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTCAAGAGAC 1067
|
|
|
QY 548 ACAGGATTTAAGTGTGCTGTGCGCAACTCTA CTGCGCTTTACATCGGT 607
|
|
|
|
Db 1068 ACAGGATTTAAGTGTGCTGTGCGCAACTCTA CTGCGCTTTACATCGGT 1127
|
|
|
QY 608 CACACTTGACATATGATTAAGCCGAGGCGAGAGAGGATGCGA 667
|
|
|
|
Db 1128 CACACTTGACATATGATTAAGCCGAGGCGAGAGAGGATGCGA 1187
|
|
|
QY 668 CTGTGTCGTCGCGAGAGAGGTGTCAAGTTTGTATGAGCATCCGTTAAGC 727
|
|
|
|
Db 1188 CTGTGTCGTCGCGAGAGAGGTGTCAAGTTTGTATGAGCATCCGTTAAGC 1247
|
|
|
QY 728 CGATTAGGCTTCATACATTGAGTAAGTCTTACATCTTTCTTTATGAG 787
|
|
|
|
Db 1248 CGATTAGGCTTCATACATTGAGTAAGTCTTACATCTTTCTTTATGAG 1307
|
|
|
QY 788 CGCATCAAGA 797
|
|
|
|
Db 1308 CGCATCAAGA 1317
|
|
|

RESULT 3
US-09-313-294A-7446

/ Sequence 7446, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Lalagudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVE.
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 7446
/ LENGTH: 283
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Inocyte ID No. 6476212 700381739H1
/ NAME/KEY: unsure
/ LOCATION: 216

/ OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-7446

Query Match 12.0%; Score 96.8; DB 3; Length 283;
Best Local Similarity 67.2%; Pred. No. 2e-21;
Matches 137; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 509 CCCAATCGATGCGGATCTTGACAGGAGCGCTGTGATGACAGATTTAAGTGTGCTGT 568
DB 2 CCAGACCGGTGGCGCAACCTGTAGGAAGCGTGTGGGTTCACGGGTTTAACTGCCGATGC 61
QY 569 GGCACCTCTACTACGCTTTTACATCGGTACTCGGACAAACACTTGCACATATGACTAC 628
DB 62 GGGACACACTACTGTTCATGACACCGCTACTCCGACAAACAGACTGCCAGTTGCACTAT 121
QY 629 AAAGCCGACGGGACAGAGACGATGCGAAAGCTTAATCTTGTGCGTGGCCGAGAGGTT 688
DB 122 CGAAGCTGACGTAGGAGCGCTATCGCCAAAGCCCAATCCAGTGTGAGAGCGGAGAAAGCTT 181
QY 689 GTCAAGTTTGTATGAGCATCCGTT 712
DB 182 GACAAGATCTGAGCGGGGCGCAT 205

RESULT 4
US-09-313-294A-1951
Sequence 1951, Application US/09313294A

/ GENERAL INFORMATION:
/ APPLICANT: Lalqudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 1951
/ LENGTH: 281
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6476212 70055176H1
US-09-313-294A-1951

Query Match 11.6%; Score 93.6; DB 3; Length 281;
Best Local Similarity 67.3%; Pred. No. 2.3e-20;
Matches 132; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 509 CCCAATCGATGCGGATCTTGACAGGAGCGTGTGATGACAGATTTAAGTGTGCTGT 568
DB 10 CCAGACCGGTGGCGGCGCTGACGAAAGCGTGTGGGTTCACGGGTTTAACTGCCGATGC 69
QY 569 GGCACCTCTACTACGCTTTTACATCGGTACTCGGACAAACACTTGCACATATGACTAC 628
DB 70 GGGACATGTACTGTTCGTGTGACCGCTACTCCGACAAACATGACTGTTCAGTTGCACTAT 129
QY 629 AAAGCCGACGGGACAGAGACGATGCGAAAGCTTAATCTTGTGCGTGGCCGAGAGGTT 688
DB 130 CGGACCTGACGACAGGAGCGGATGCGCAAGCCCAATCTCTGTGTGAGGCGGAGAAAGCTC 189
QY 689 GTCAAGTTTGTATGAG 704
DB 190 GACAAGATCTGAGGCG 205

RESULT 5
US-09-313-294A-3856
Sequence 3856, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:

/ APPLICANT: Lalqudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVE
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 3856
/ LENGTH: 211
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700282265H2
/ NAME/KEY: unsure
/ LOCATION: 41, 60, 99, 116, 118, 135, 154
/ OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3856

Query Match 9.8%; Score 78.8; DB 3; Length 21
Best Local Similarity 64.5%; Pred. No. 1.4e-15;
Matches 127; Conservative 0; Mismatches 69; Indels

QY 509 CCCAATCGATGCGGATCTTGACAGGAGCGTGTGATGACAGATTTA 568
DB 9 CCAGACCGGTGGCGGCGCTGACGAAAGCGTGTGGGTTCACGGGATTTA 68
QY 569 GGCACCTCTACTACGCTTTTACATCGGTACTCGGACAAACACTTGC 627
DB 69 GGGACATGTACTGTTCGTGTGACCGCTACTCCGACAAACAGTACTAT 128
QY 628 CAAGCCGACGGGACAGAGACGATGCGAAAGCTTAATCTTGTGCTG 687
DB 129 TCGACATGACGACAGAGAGCGGATTTCCAGGCGCAATCTGTGTGAGG 188
QY 688 TGTCAAGTTTGTATGAG 704
DB 189 GACAAGATCTGAGGCG 205

RESULT 6
US-08-861-269-4
Sequence 4, Application US/08861269
Patent No. 5817494

/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Corley, Neil C.
/ APPLICANT: Lal, Preeti
/ TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/861,269
/ FILING DATE: Filed Herewith
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-08-861-269-4

Query Match 9.6%; Score 77; DB 2; Length 1197;
Best Local Similarity 63.0%; Pred. No. 1.4e-14;
Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 512 AATGATCGGATCTTGACAGAAAGCGTGTGATTGACAGATTAAAGTTCGCTGGC 571
DB 700 AATGCTGTTTCATGTGACGAGAAAGTGAGACTTACTGGTTGAATGCCGGTGTGA 759
QY 572 AACTCTACTGCGCTTTACATCGGTACTCGACAAACACTTGACATATGACTACAA 631
DB 760 AATGTTACTGTGTGTACACCGTTACTCAGATGTACACAAATGCTTTACATTAACAA 819
QY 632 GCCGACGGCAGAAAGCGATTGCGAAAGCTAATCTTGTGTGCGCCGAGAGGTGTC 691
DB 820 GCCGATGCTGTGAGAAATCAGAAAAGAAATCCAGTAGTGTGTGAAAAAGATCCAA 879
QY 692 AAGTTTGA 700
DB 880 AAGATTGA 888

RESULT 7

US-09-134-596-4
Sequence 4, Application US/09134596
Patent No. 592318
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,596
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/861,269
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-134-596-4

Query Match 9.6%; Score 77; DB 2; Length 1197;
Best Local Similarity 63.0%; Pred. No. 1.4e-14;
Matches 119; Conservative 0; Mismatches 70; Indels 0;

QY 512 AATGATCGGATCTTGACAGAAAGCGTGTGATTGACAGATTAAAGTTC 571
DB 700 AATGCTGTTTCATGTGACGAGAAAGTGAGACTTACTGGTTGAATGCCGGTGTGA 759
QY 572 AACTCTACTGCGCTTTACATCGGTACTCGACAAACACTTGACAT 631
DB 760 AATGTTACTGTGTGTACACCGTTACTCAGATGTACACAAATGCTTTACATTAACAA 819
QY 632 GCCGACGGCAGAAAGCGATTGCGAAAGCTAATCTTGTGTGCGCCG 691
DB 820 GCCGATGCTGTGAGAAATCAGAAAAGAAATCCAGTAGTGTGTGAAAAAGATCCAA 879
QY 692 AAGTTTGA 700
DB 880 AAGATTGA 888

RESULT 8

US-09-293-273-4
Sequence 4, Application US/09293273
Patent No. 6057112
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/293,273
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,596
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-855-0555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-293-273-4

Query Match
Best Local Similarity 9.2%; Score 77; DB 3; Length 1197;
Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

512 ATTCGATGCGGATCTTGCAGAGAGCGTGTGATGATGAGATTTAAGTGTGCTGTGC 571
|||||
700 ATTCGATGCTTCATGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 759
|||||
572 AACCTTACTGCGCTTACATGCGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 631
|||||
760 AATGTTACTGCTGTGTACACCGTTACTGATGATGATGATGATGATGATGATGATGATG 819
|||||
632 GCGGAGGCGAG 691
|||||
820 GCGGATGCTGTGAG 879
|||||
692 AAGTTTGA 700
|||||
880 AAGATTGA 888

RESULT 9
US-09-949-016-810
Sequence 810, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 810
LENGTH: 2425
TYPE: DNA
ORGANISM: Human
US-09-949-016-810

Query Match
Best Local Similarity 9.2%; Score 74.2; DB 3; Length 2425;
Matches 118; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

498 CTAGACCCGAGCCCATTCGATGCGATCTTGCAGAGAGCGTGTGATGAGAGATTGA 557
|||||
728 CCAACCAAG 787
|||||
558 AGTGTGCTGTGCAACCTTACTGCGCTTACATGCGTGTGATGAGAGAGAGAGAGAGAGAG 617
|||||
788 ACTGCCGATGAG 847
|||||
618 CATATGACTACAAAGCGGAG 677
|||||
848 CCGTATGATTAACAAG 907
|||||
678 CCGAGAGGTT 688
|||||
908 CTGAAAAAATT 918

RESULT 10
US-09-949-016-3693
Sequence 3693, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 3693
LENGTH: 2684
TYPE: DNA
ORGANISM: Human
US-09-949-016-3693

Query Match
Best Local Similarity 9.2%; Score 74.2; DB 3; Length 26
Matches 118; Conservative 0; Mismatches 73; Indels 0;

498 CTAGACCCGAGCCCATTCGATGCGATCTTGCAGAGAGCGTGTGATTT 557
|||||
987 CCAACCAAG 1046
|||||
558 AGTGTGCTGTGCAACCTTACTGCGCTTACATGCGTGTGATGAGAGAGAGAGAGAGAGAG 617
|||||
1047 ACTGCCGATGAG 1106
|||||
618 CATATGACTACAAAGCGGAG 677
|||||
1107 CCGTATGATTAACAAG 1166
|||||
678 CCGAGAGGTT 688
|||||
1167 CTGAAAAAATT 1177

RESULT 11
US-09-244-805-10
Sequence 10, Application US/09244805
Patent No. 6699660

GENERAL INFORMATION:
APPLICANT: Worley, Paul F.
APPLICANT: Lanahan, Anthony
APPLICANT: Goetz, Bernard
APPLICANT: Heinisch, Holger
APPLICANT: Kune, Rohini
APPLICANT: Scheek, Sigrid
APPLICANT: Nikolich, Karoly
APPLICANT: Zhukovski, Eugene
TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE;
FILE REFERENCE: 10496/004001
CURRENT APPLICATION NUMBER: US/09/244,805
CURRENT FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/074,518
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: 60/074,135
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 2515
TYPE: DNA

ORGANISM: Eukaryote
FEATURE:
NAME/KEY: CDS
LOCATION: (414)...(1055)
US-09-244-805-10

Query Match
Best Local Similarity 8.9%; Score 71.8; DB 3; Length 2515;
Matches 115; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 498 CTAGACCCGACCCCATCATCGATCCGATCTTTCAGAGAAAGCGTTGGATTGACAGATTGA 557
DB 853 CCAAAACCAAGAAAGAACAGATGTTTATGTGTAGAAAGAAAGTTGGCCCTTACAGCGTTTG 912
QY 558 AGTGCCTGTGTGGCAACCTCTACTGCGCTTTCATCTGCTACGAGCAAAACGACCTTGCA 617
DB 913 ACTGCCGAGTGTGAATTTGTTTGTGACCTTCACCGTTACTCTACAGCAACACTGTC 972
QY 618 CATATGACTACAAAGCCGACGAGGAGCAAGCGATTGCGAAAGCTATCCTCTTGTGTTG 677
DB 973 CTTATGATTACAAAGACAGAAAGCTGCAGCAAAAATCAGAAAATAATCCAGTTGTTGCG 1032
QY 678 CCGAGAA 684
DB 1033 CTGAAAA 1039

RESULT 12
US-09-949-016-12552
Sequence 12552, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12552
LENGTH: 14524
TYPE: DNA
ORGANISM: Human
US-09-949-016-12552

Query Match
Best Local Similarity 7.6%; Score 61; DB 3; Length 14524;
Matches 91; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 548 ACAGATTTAAGTGTGCTGTGGCAACCTCTACTGCGCTTTCATCTGCTACGAGCAAA 607
DB 10877 ACAGGTTTGAATGCGAGTGTGAATTTGTTTGTGACCTTCACCGTTACTACAAAG 10936
QY 608 CACACTTCACATATGATCTACAAAGCCGACGAGGAGCAAGCGATTGCGAAAGCTATCT 667
DB 10937 CACAACTGCTGTAATGATTACAAAGACAGAAAGCTGCAGCAAAAATCAGAAAAGAAATCCA 10996
QY 668 CTTGTCTGTGGCCGAGAAAGTT 688
DB 10997 GTTGTGTGTGCTGAAAAAATT 11017

RESULT 13
US-09-949-016-15435
Sequence 15435, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15435
LENGTH: 14785
TYPE: DNA
ORGANISM: Human
US-09-949-016-15435

Query Match
Best Local Similarity 7.6%; Score 61; DB 3; Length 1478
Matches 91; Conservative 0; Mismatches 50; Indels 0;

QY 548 ACAGATTTAAGTGTGCTGTGGCAACCTCTACTGCGCTTTCATCTGCGT 607
DB 11137 ACAGGTTTGAATGCGAGTGTGAATTTGTTTGTGACTTCACCGTT 11196
QY 608 CACACTTCACATATGATCTACAAAGCCGACGAGGAGCAAGCGATTGCGA 667
DB 11197 CACAACTGCTGTAATGATTACAAAGACAGAAAGCTGCAGCAAAAATCAGAA 11256
QY 668 CTTGTCTGTGGCCGAGAAAGTT 688
DB 11257 GTTGTGTGTGCTGAAAAAATT 11277

RESULT 14
US-09-640-211A-1742
Sequence 1742, Application US/09640211A
Patent No. 6833446
GENERAL INFORMATION:
APPLICANT: Wood, Marion
APPLICANT: Shenk, Michael A.
APPLICANT: McGrath, Annekte
APPLICANT: Glenn, Matthew
TITLE OF INVENTION: Compositions and Methods for the
FILE REFERENCE: Modification of Gene Transcription
CURRENT APPLICATION NUMBER: US/09/640,211A
CURRENT FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 2368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1742
LENGTH: 348
TYPE: DNA
ORGANISM: Pinus radiata
US-09-640-211A-1742

Query Match
Best Local Similarity 7.4%; Score 59.4; DB 3; Length 34
Matches 90; Conservative 0; Mismatches 51; Indels 0;

QY 512 AATGATCCGATCTTTCAGAGAACGTTGTGATTGACAGATTAAAGT 571
DB 204 AAGAGTGTTTGTGCTGTGAACAGCGGTTGGCGCTTAAAGT 263
QY 572 AACCTTACTGCGCTTTCATCTGCTACGAGCAAAACACACTTGACAT 631
DB 264 GACCTTTTCTGCGCTTCAGCAAGGTAATCTGATATGATGACTCTCTT 323
QY 632 GCCGAGGAGCAAGAAAGATT 652

Db 324 ACTGCCGCCGCTCGGCATT 344

```

RESULT 15
US-09-313-294A-4836
; Sequence 4836, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laljudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4836
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700349114H1
; NAME/KEY: unsure
; LOCATION: 18, 21
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4836
    
```

```

Query Match
Best Local Similarity 61.8%; Score 54.6; DB 3; Length 277;
Matches 87; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 156 AGAGACGACCTCGCAGGCCCTGAGGGTTCAGATTATGTCAAGAACCTTTGGCGCTTCT 215
DB 122 AGGAGCTGGGTGCGCAGGCCGCCGAGGAGCCATCTCTGCAATCAATACTGTGGCTTCT 181
QY 216 TGGCAGCCCAAGTACATGAGGGTTGTCTCGAAGTCTTACCGAGACAGTCATGCAAG 275
DB 183 TGGCAGCCCGCGGCGACATGATGCTCCAGAGTCCCAAGAGAGATGATTAACGAGC 241
QY 276 CGAAGTACGGGCTTACGCTG 296
DB 242 AGGATCAGGCCCAAGCTGCTG 262
    
```

Search completed: December 8, 2005, 10:15:34
 Job time : 200 secs

This Page Blank (uspto)

Db 241 GTGCTGAAGTGTCTACCGAGAGACATCATGCAAGATGACGAGCTTTAGCTGACGA 300
Qy 301 AGCCACTAAGGCTGTCTACGGCGGACATCTGSCCAACCTGTGTCTTCAGCCCCCGCTCC 360
Db 301 AGCCACTAAGGCTGTCTACGGCGGACATCTGSCCAACCTGTGTCTTCAGCCCCCGCTCC 360
Qy 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAAACAATGTTGTGCCCATCAATC 420
Db 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAAACAATGTTGTGCCCATCAATC 420
Qy 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGACGTGCTCCCTCAGGCAATGATC 480
Db 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGACGTGCTCCCTCAGGCAATGATC 480
Qy 481 CTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAGGAAAGCGTGT 540
Db 481 CTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAGGAAAGCGTGT 540
Qy 541 TGGATTGACAGGATTTAAAGTGTGCTGTGGCAACCTCTACTGCGTTTACATCGGTACTC 600
Db 541 TGGATTGACAGGATTTAAAGTGTGCTGTGGCAACCTCTACTGCGTTTACATCGGTACTC 600
Qy 601 GGACAAACACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCGAAAGC 660
Db 601 GGACAAACACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCGAAAGC 660
Qy 661 TTAATCTCTGTGCGGCGCGAGAGGTTGTCAAGTTTGTATGATGAGATCCGTTAACTTTT 720
Db 661 TTAATCTCTGTGCGGCGCGAGAGGTTGTCAAGTTTGTATGATGAGATCCGTTAACTTTT 720
Qy 721 CTGCCGAGATTTAGGCTTTCATACATTTGAGTAACCTTCTTCTTTATCGAAG 780
Db 721 CTGCCGAGATTTAGGCTTTCATACATTTGAGTAACCTTCTTCTTTATCGAAG 780
Qy 781 AGCGAGTGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTGCATCAAGAGCTCGCC 804

RESULT 2
US-10-716-089-12
; Sequence 12, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSMALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/10/716,089
; CURRENT FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: 60/196,001
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 12
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Physcomitrella patens
US-10-716-089-12

Query Match 100.0%; Score 804; DB 7; Length 804;
Best Local Similarity 100.0%; Pred. No. 1,2e-265; Indels 0; Gaps 0;
Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCCGGGACAGTCCGCTTAAAGTGTGTCATTAAGTGTGTCGAAGTGAAGC 60
Db 1 ATCCGGGACAGTCCGCTTAAAGTGTGTCATTAAGTGTGTCGAAGTGAAGC 60
Qy 61 CTGAGCGAGATTTGCAAGATTTTCTCATACGCTTCTGATTGAAAGATACACCTTAT 120
Db 61 CTGAGCGAGATTTGCAAGATTTTCTCATACGCTTCTGATTGAAAGATACACCTTAT 120

Db 61 CTGAGCGAGATTTGCAAGATTTTCTCATACGCTTCTGATTGAAAGAT 120
Qy 121 TAGTCTGTTAAAGATGAGCCAGAGCGGTGTCTCAGAGAGCGACTGCG 180
Db 121 TAGTCTGTTAAAGATGAGCCAGAGCGGTGTCTCAGAGAGCGACTGCG 180
Qy 181 GGGTCCAGTTATGTGCAAGAACCTTTGCGGCTTCTTCGCGACCAAGCT 240
Db 181 GGGTCCAGTTATGTGCAAGAACCTTTGCGGCTTCTTCGCGACCAAGCT 240
Qy 241 GTGCTGAAGTGTCTACCGAGACATGATATGCAAGCGCAAGATGACGCT 300
Db 241 GTGCTGAAGTGTCTACCGAGACATGATATGCAAGCGCAAGATGACGCT 300
Qy 301 AGCCACTAAGGCTGTCTACGGCGGACATCTGSCCAACCTGTGTCTTCAG 360
Db 301 AGCCACTAAGGCTGTCTACGGCGGACATCTGSCCAACCTGTGTCTTCAG 360
Qy 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAAACAATGATTTGTG 420
Db 361 TGTACATGAGACCAAGCTCACAATGCGAGGTTGAGAAACAATGATTTGTG 420
Qy 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGACGTGCTCCCTCAG 480
Db 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGACGTGCTCCCTCAG 480
Qy 481 CTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATGATGCGGATTTTTC 540
Db 481 CTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATGATGCGGATTTTTC 540
Qy 541 TGGATTGACAGGATTTAAAGTGTGCTGTGGCAACCTCTACTGCGTTTAA 600
Db 541 TGGATTGACAGGATTTAAAGTGTGCTGTGGCAACCTCTACTGCGTTTAA 600
Qy 601 GGACAAACACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCG 660
Db 601 GGACAAACACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCG 660
Qy 661 TTAATCTCTGTGCGGCGCGAGAGGTTGTCAAGTTTGTATGATGAGCATCC 720
Db 661 TTAATCTCTGTGCGGCGCGAGAGGTTGTCAAGTTTGTATGATGAGCATCC 720
Qy 721 CTGCCGAGATTTAGGCTTTCATACATTTGAGTAACCTTCTTCTTTATCG 780
Db 721 CTGCCGAGATTTAGGCTTTCATACATTTGAGTAACCTTCTTCTTTATCG 780
Qy 781 AGCGAGTGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTGCATCAAGAGCTCGCC 804

RESULT 3
US-09-828-303-4
; Sequence 4, Application US/09828303
; Patent No. US20020102695A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSMALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROT
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Physcomitrella patens

US-09-828-303-4

Query Match 96.6%; Score 776.4; DB 3; Length 1531;
Best Local Similarity 99.7%; Pred. No. 5.4e-256;
Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 8 GCACCAAGTCCCGCTTAGTGTGTGTCATTTAGTGTGTGCAAGCTGAAGCTTGAGC 67
DB 529 GCACCAAGTCCCGCTTAGTGTGTGTCATTTAGTGTGTGCAAGCTTGAGC 588
DB 68 GAGATTGCAAGATTCTCATACGCTTCTGATTAGAAAAGATACCCCTTATTAGTCTG 127
DB 589 GAGATTGCAAGATTCTCATACGCTTCTGATTAGAAAAGATACCTTATTAGTCTG 648
QY 128 TTAAGATGCGCACCGAGCGGTGTCTCAGSAGACGACCTCCAGAGCCCTTGAGGTCCA 187
DB 649 TTAAGATGCGCACCGAGCGGTGTCTCAGSAGACGACCTCCAGAGCCCTTGAGGTCCA 708
QY 188 GTTATGTGCAAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATGSGGTTGTGCTCG 247
DB 709 GTTATGTGCAAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATGSGGTTGTGCTCG 768
QY 248 AAGTCTACCGAGAGACAGTCAATGCAAGGGAAGATGACGGCTTTAGCTGAGCAAGCCACT 307
DB 769 AAGTCTACCGAGAGACAGTCAATGCAAGGGAAGATGACGGCTTTAGCTGAGCAAGCCACT 827
QY 308 CAGGCTGCTCAGGCGACATCTGCGCAAGCTGCTGCTTCAAGCCCTCCGCTCTGTACT 367
DB 828 CAGGCTGCTCAGGCGACATCTGCGCAAGCTGCTGCTTCAAGCCCTCCGCTCTGTACT 887
QY 368 GAGACCAAGCTCACAATGCGAGGTTGAGAAACATGATGTGCGCCATCACTTCCAGC 427
DB 888 GAGACCAAGCTCACAATGCGAGGTTGAGAAACATGATGTGCGCCATCACTTCCAGC 947
QY 428 TATCAACAAGACTGTGTTACCCCGCTGACGCTGCTCAGSAGATGAAATCTCTATC 487
DB 948 TATCAACAAGACTGTGTTACCCCGCTGACGCTGCTCAGSAGATGAAATCTCTATC 1007
QY 488 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAGAGAAAGCTGTGAGATTG 547
DB 1008 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAGAGAAAGCTGTGAGATTG 1067
QY 548 ACAGGATTTAAGTGTGCGTGTGCGCAACCTCTACTCGCTTTACATCGGATCTCGACAA 607
DB 1068 ACAGGATTTAAGTGTGCGTGTGCGCAACCTCTACTCGCTTTACATCGGATCTCGACAA 1127
QY 608 CACACTTGACATATGACTACAAAGCCGACGAGGAGAAAGCATTCGAAAGCTATCT 667
DB 1128 CACACTTGACATATGACTACAAAGCCGACGAGGAGAAAGCATTCGAAAGCTATCT 1187
QY 668 CTGTGCGTGGCGGAGAAAGTGTCAAGTTTGTATGAGCATCCGTTAAGCTTTTCTGCCGA 727
DB 1188 CTGTGCGTGGCGGAGAAAGTGTCAAGTTTGTATGAGCATCCGTTAAGCTTTTCTGCCGA 1247
QY 728 CGATTAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
DB 1248 CGATTAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1307
QY 788 CGCATCAAGA 797
DB 1308 CGCATCAAGA 1317

RESULT 4

US-10-716-089-4

Sequence 4, Application US/10716089

Publication No. US20040107463A1

GENERAL INFORMATION:

APPLICANT: COSTA E SILVA, OSWALDO DA

APPLICANT: BOHNER, HANS J.

APPLICANT: VAN THIELEN, NOCHA

APPLICANT: CHEN, ROUING

TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND

TITLE OF INVENTION: METHODS OF USE IN PLANTS

FILE REFERENCE: 16313-0030

CURRENT APPLICATION NUMBER: US/10/716,089

CURRENT FILING DATE: 2003-11-18

PRIOR APPLICATION NUMBER: 60/196,001

PRIOR FILING DATE: 2000-04-07

NUMBER OF SEQ ID NOS: 79

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 4

LENGTH: 1531

TYPE: DNA

ORGANISM: Physcomitrella patens

US-10-716-089-4

Query Match 96.6%; Score 776.4; DB 7; Length 1

Best Local Similarity 99.7%; Pred. No. 5.4e-256;

Matches 788; Conservative 0; Mismatches 1; Indels 1;

QY 8 GCACCAAGTCCCGCTTAGTGTGTGTCATTTAGTGTGTGCAAGCTTGAGC 67
DB 529 GCACCAAGTCCCGCTTAGTGTGTGTCATTTAGTGTGTGCAAGCTTGAGC 588
QY 68 GAGATTGCAAGATTCTCATACGCTTCTGATTAGAAAAGATACACC 127
DB 589 GAGATTGCAAGATTCTCATACGCTTCTGATTAGAAAAGATACATC 648
QY 128 TTAAGATGCGCACCGAGCGGTGTCTCAGSAGACGACCTCCAGAGCCCT 187
DB 649 TTAAGATGCGCACCGAGCGGTGTCTCAGSAGACGACCTCCAGAGCCCT 708
QY 188 GTTATGTGCAAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATGSGGTTGTGCTCG 247
DB 709 GTTATGTGCAAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATGSGGTTGTGCTCG 768
QY 248 AAGTCTACCGAGAGACAGTCAATGCAAGGGAAGATGACGGCTTTAGCTG 307
DB 769 AAGTCTACCGAGAGACAGTCAATGCAAGGGAAGATGACGGCTTTAGCTG 827
QY 308 CAGGCTGCTCAGGCGACATCTGCGCAAGCTGCTGCTTCAAGCCCTCCGCTCTGTACT 367
DB 828 CAGGCTGCTCAGGCGACATCTGCGCAAGCTGCTGCTTCAAGCCCTCCGCTCTGTACT 887
QY 368 GAGACCAAGCTCACAATGCGAGGTTGAGAAACATGATGTGCGCCATCACTTCCAGC 427
DB 888 GAGACCAAGCTCACAATGCGAGGTTGAGAAACATGATGTGCGCCATCACTTCCAGC 947
QY 428 TATCAACAAGACTGTGTTACCCCGCTGACGCTGCTCAGSAGATGAAATCTCTATC 487
DB 948 TATCAACAAGACTGTGTTACCCCGCTGACGCTGCTCAGSAGATGAAATCTCTATC 1007
QY 488 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAGAGAAAGCTGTGAGATTG 547
DB 1008 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAGAGAAAGCTGTGAGATTG 1067
QY 548 ACAGGATTTAAGTGTGCGTGTGCGCAACCTCTACTCGCTTTACATCGGATCTCGACAA 607
DB 1068 ACAGGATTTAAGTGTGCGTGTGCGCAACCTCTACTCGCTTTACATCGGATCTCGACAA 1127
QY 608 CACACTTGACATATGACTACAAAGCCGACGAGGAGAAAGCATTCGAAAGCTATCT 667
DB 1128 CACACTTGACATATGACTACAAAGCCGACGAGGAGAAAGCATTCGAAAGCTATCT 1187
QY 668 CTGTGCGTGGCGGAGAAAGTGTCAAGTTTGTATGAGCATCCGTTAAGCTTTTCTGCCGA 727
DB 1188 CTGTGCGTGGCGGAGAAAGTGTCAAGTTTGTATGAGCATCCGTTAAGCTTTTCTGCCGA 1247
QY 728 CGATTAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
DB 1248 CGATTAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1307
QY 788 CGCATCAAGA 797
DB 1308 CGCATCAAGA 1317

RESULT 5
US-10-424-599-99368
; Sequence 99368, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 99368
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_60743C.1
US-10-424-599-99368
Query Match 13.3%; Score 106.6; DB 7; Length 1385;
Beet Local Similarity 68.2%; Pred. No. 3.8e-25;
Matches 148; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 509 CCACATCGATGCGATCTTTCAGAGAGCGTGTGGATTGACAGATTAAAGTGTGCTGT 568
DB 805 CCAATTCATGAGAGCACTGCAATTAAGCGGTTGGTTGACAGATTAAATTTGTGATGT 864
QY 569 GGCACCTCTACTGCGCTTTTACATCGTACTCGGCAACAACACACTTGCACATATACATAC 628
DB 865 GGTAACTTTCTGTCTGTATCATCTGCTACTCAGCAAGCATTAATTGCCATTGATTAC 924
QY 629 AAAGCGGAGGCGAGAGCGATTCGAAAGCTAATCCCTTGTGTCGCGGCGAGAGGT 688
DB 925 CACACTGCTGCGAGAGATCCATAGCTAAGCGAATTCGCGCTGTCAAGTTGAGAACTT 984
QY 689 GTCAAGTTTGTATGAGCATCCGTTAAGCTTTTCTGCC 725
DB 985 GATAAGATCTAATTTGTTGTCTTTAAGTTTCATGCC 1021
RESULT 6
US-10-425-114-22561
; Sequence 22561, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22561
; LENGTH: 1003
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3587-277-F2_FLI
US-10-425-114-22561
Query Match 12.9%; Score 103.4; DB 7; Length 1003;
Beet Local Similarity 53.6%; Pred. No. 4.1e-24;
Matches 302; Conservative 0; Mismatches 231; Indels 3;

Matches 302; Conservative 0; Mismatches 231; Indels 3;
QY 156 AGGAGAGCACTCCAGAGCCCTTGAAGTCCAGTTATGTGAAGACCTT 215
DB 122 AGGAGGCTGGGTGCCAGGCCCCCGAGGAGACCATCTCTGATCAATATAC 181
QY 216 TCGGACGCAAGTACATGAGGAGTGTGCTGGAAGTGTCAACGAGAGAT 275
DB 182 TCGGACGCGCGGCGACATGAAATGTGCTCCAAAGTGCACAAGAGAT 241
QY 276 CGAAGATGACGCGTTTGAAGTGAAGACCACTCAGGCTCTCAGCGCAT 335
DB 242 AGGATCAGGCCCAAGCTGCTGC-----CTCCTCTATCGAC 287
QY 336 CTGCTGCTGTTCAGCCCCCGCTCCTGATGATGAGACCAACTACAG 395
DB 288 CGGACGACGCGCT-----CATGAGCCGGTTGTGCTG 333
QY 396 GAACATGATTTGTGCGGATCAATCTTCAGCTATCAACAAGACCTGT 455
DB 334 GTAGTAGCTGTGGCCAAAGTTGAGTTGCAACATGAAAC--GTGAGCA 391
QY 456 CAGTGCCTTCAGGCAAGTGAATCTCTATGCAAGCTCCCTTAGACC 515
DB 392 TTGCCCGAACCCAGCGAGGGGTGGCGGCGATCTCCAAAGGGGAAAGGT 451
QY 516 GATCGCATTTTGAAGAGAGGTGTGGATTGACAGATTAAAGTGTG 575
DB 452 GGTGAGGCGCTTCGAGAGAGGAGGTGACTTACGGGATTCACATGCGG 511
QY 576 TCTACTGCGCTTTACATCGTACTCGACAAACACACTTGCACATATGA 635
DB 512 TGTACTGCGCACTCCACGCTACTCCGACAAAGCAAGATGCAAGTTGCA 571
QY 636 CAGGCGAGAGAGCATTCGCAAGCTAATCTCTTGTGTCGCGAGAA 695
DB 572 CTGCCAGGAGCGCATTCGCAAGCTAATCCGTGTGTGAAGGACAGAA 631
QY 696 TTTGATGAGCATCCGTTAAGCTT 718
DB 632 TCTAGGGGGGTTCTTTACGGTT 654
RESULT 7
US-10-425-114-27861
; Sequence 27861, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27861
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4740-074-B6_FLI
US-10-425-114-27861
Query Match 12.8%; Score 103; DB 7; Length 873
Beet Local Similarity 53.2%; Pred. No. 5.2e-24;
Matches 307; Conservative 0; Mismatches 240; Indels 3;
QY 156 AGGAGACGACTCCAGAGCCCTTGAAGTCCAGTTATGTGAAGACCTT 215

```

Db      144 AGGAGCTGGGTGCGAGGCCCCCGAGGAGCCATCTCTGCACTCAATTAAGTGGCTTCT 203
Qy      216 TCGGCGCCAGCAAGTACCATGGGGTTGTCTCGAAGTCTACCGAGAGACATGATCAAG 275
Db      204 TCGGCGCCGCGCGGACCATGAAACATGTGCTTAAGTCCCAAGAGAGATGATTAACGAAG 263
Qy      276 CGAAGATGACGGCTTAAAGTGAAGCAAGCACTAGGCTCTGAGGAGCAATCTGCCACAG 335
Db      264 AGGATCAGGCGCAAGCTGGCTGC-----CTCCTTAATGACAGCATCGTGA 309
Qy      336 CTGCTGCTGTTGAGCCCCCGCTCTGTATCATGAGACCAAGCTCACATGCGAGTTGAGA 395
Db      310 CGGCGAGCGAGCGCT-----CATGAGCCGGTTGTGCTGCGACCAACAG 355
Qy      336 GAACAATGATTTGGCCGATCAATCTTCAAGCTATCAACAAGCCTGGTTACCCCGCTG 455
Db      356 GTAGTAGCTGTGTGCCCAAGTCCAGTTGCAAAACATGAC--GTGCGAGCGCCGCTGATG 413
Qy      456 CAGCTGCCCTCAGGAGTAGAAGTCTCTATGCGCAGCTCCCTTGAACCCGAGCCCAATC 515
Db      414 TTGCGGAGCCGACGAGGGGGGTGGCGGCGATCTCCAAAGGGGAGGTAGGGCCGAAAC 473
Qy      516 GATGCGGATCTTTCAGAGAACGCTGTGATGACAGATTTTAAGTGTGCTGTGGCAAC 575
Db      474 GGTGACGCGCTTCGAGAGAGAGGTTGGCTCAGCGGATTTCAACTGCCGTGGAACT 533
Qy      576 TCTACTGCGCTTAACTGCTGTACTGGAACAACACTTGCACATATGACTAACAAAGCCG 635
Db      534 TGTACTGCGCGCTTCACCGCTACTCCGACAGACGACTGCAAGTTGCACTCCGGAAC 593
Qy      636 CAGGCGAGAAAGCAATTTGGAAAGCTAATCTTGTGTGCGCCGAGAAAGTTGTCAAGT 695
Db      594 CTGCCAGGAGCGCATGTCAGAGGCTAATCCGGTGTGAAGGCAAGCAAGCTGCACAGA 653
Qy      696 TTTGATGAGCATCCGTTAACTTTTCTGCGCAGCAAT 732
Db      654 TCTAGGGGGGGTTCTTACGGTTGTGTCTGAGAAAGATT 690

```

```

RESULT 8
US-10-425-114-26146
; Sequence 26146, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 26146
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LJB4073-024-B8_FLI
US-10-425-114-26146

```

```

Query Match      12.8% Score 103; DB 7; Length 945;
Best Local Similarity 52.3%; Pred. No. 5.4e-24;
Matches 284; Conservative 0; Mismatches 220; Indels 39; Gaps 1;
Qy      156 AGAGACGACCTCGAGGCGCCCTGAGGCTCAGTTATGTGCAAGAACCTTTGGCGCTTCT 215
Db      113 AGGAGACTGAGATGCGAGGACCTGAGGAGCCATCATCTTTGATCAATATAGCGGCTTCT 172

```

```

Qy      216 TCGGCGCCAGCAAGTACCATGGGGTTGTCTCGAAGTCTACCGAGAGAC 275
Db      173 TCGGCGCGGCGAGCTACCAAGAACATGTGCTCCAAAGTCCCAAGAGAGA 232
Qy      276 CGAAGATGACGGCTTAAAGTGAAGCAAGCACTAGGCTCTGAGGAGAC 335
Db      233 AGGACAGGCGCAAGCTGGCTG----- 253
Qy      336 CTGCTGCTGTTGAGCCCCCGCTCTGTATCATGAGACCAAGCTCAATC 395
Db      254 CCTCTTAATGACAGCATCTGCAATGGCAAGATGTCTGATGAAAC 313
Qy      396 GAACAATGATTTGGCCGATCAATCTTCAAGCTATCAACAACCTGCT 455
Db      314 GCAACAGAGTGTGCTGTGCTTCCAAATGAGTTCAACAATGAAAGT 373
Qy      456 CAGTGCCCTCAGGAGTAGAAGTCTCTATGCGAGCTCCCTTGAAC 515
Db      374 ATGTTGCTGAGACTTACGAGAGGGGGCGCGGTGATCTCCAAAGGAAAGT 433
Qy      516 GATGCGGATCTTTCAGAGAACGCTGTGATGACAGATTTTAAGTGTG 575
Db      434 GGTGACGACCTTTCAGAGAAAGAGGTTGACTTACAGATTCACACTGCC 493
Qy      576 TCTACTGCGCTTAACTGCTGTACTGGAACAACACTTGCACATATG 635
Db      494 TGTACTGTGCACTGACCGCTACTCCGACAGACGACTGCAAGTTG 553
Qy      636 CAGGCGAGAAAGCAATTTGGAAAGCTAATCTTGTGTGCGCGAGAG 695
Db      554 CTGCTAGGAGATGCCATTGCCAAGGCTAATCCAGTGTGAAGCGGACAA 613
Qy      696 TTT 698
Db      614 TCT 616

```

```

RESULT 9
US-10-425-115-71431
; Sequence 71431, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 71431
; LENGTH: 947
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: WRT4577_16514C.1
US-10-425-115-71431

```

```

Query Match      12.8% Score 103; DB 8; Length 947
Best Local Similarity 52.3%; Pred. No. 5.4e-24;
Matches 284; Conservative 0; Mismatches 220; Indels 1;
Qy      156 AGAGACGACCTCGAGGCGCCCTGAGGCTCAGTTATGTGCAAGAACCT 215
Db      115 AGGAGACTGAGATGCGAGGACCTGAGGAGCCATCTTTGATCAATTA 174
Qy      216 TCGGCGCCAGCAAGTACCATGGGGTTGTCTCGAAGTGTCTACCGAGAG 275
Db      175 TCGGCGAGCGAGCTTACCATGAAACATGTGCTCCAGTGTCCAAAGAGAT 234
Qy      276 CGAAGATGACGGCTTAAAGTGAAGCAAGCACTAGGCTCTCAGGCGAC 335

```

```
Db      225 AGGAGCAGGCCAAGTGGCTG----- 255
Qy      336 CTGCTGCTGTTGAGCCCCCGCTCTCTGATCATGAGCAAGCTCATGCGAGTTGAGA 395
Db      256 CCTCTCTATGACAGCATGCTCAATGCGACGATGCTGTATGAGCAACAGTTGTTGCTG 315
Qy      396 GAACAATGATTGCGCGATCAATTTCCAGCTATCAACAAGACTGGTTACCCCGCTG 455
Db      316 GCACACAGTGGTGGCTGCTGCTCTCAATGAGTTGCAACAATGAACCTGACGCCGCTG 375
Qy      456 CAGTCCCTCAGCAGTGAAGTCTCTATGCAAGTCCCTCTAGACCCGACCCCAATC 515
Db      376 ATGTTGCTGAGACTGACGAGGGGGGGGGGTATCTCCAAAGGAAAGTGGGCGGAAC 435
Qy      516 GATGCGGATCTTGGCAGAAAGCGTTGGATTGACAGATTAAAGTGTGCTGGCAAC 575
Db      436 GGTGACGACTTGGCAGAAAGGTTGGACTTACAGATTCAACTGCCGCTGTGGAACT 495
Qy      576 TCTACTGGGCTTTACATGGTACTCGGACAAACACACTTGCATATGACTACAAAGCG 635
Db      496 TGTACTGTGACTGACCGCTACTCCGACACGACGACTGCAAGTTCCATATGCGACTG 555
Qy      636 CAGGCGAGGAAGCGATTGGGAAAGCTAACTCTTGTGCTGGCCGAGAAAGTTGCAAGT 695
Db      556 CTGCTAGGGATGCCATTGGCAGAGGCTAATCCAGTGTGAAGCGGACAAAGCTCGACAA 615
Qy      696 TTT 698
Db      616 TCT 618
```

```
RESULT 10
US-10-767-701-15574
; Sequence 15574, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 15574
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS3087_1
US-10-767-701-15574
```

```
Query Match      12.8%; Score 103; DB 7; Length 1142;
Best Local Similarity 52.3%; Pred. No. 66-24;
Matches 284; Conservative 0; Mismatches 220; Indels 39; Gaps 1;

Qy      156 AGGAGCAGCTCGCAGGCCCTGAGGGTCCAGTTATGTGCAAGAACTTTGCGGCTTCT 215
Db      173 AGGAGCTGATGACGACCTGAGGGACCATCTTTGCATCAATAACTCGGCTTCT 232
Qy      216 TGGGACGCCAAGCTACATGAGGGTGTGCTGCAAGTCTACCGAAGACAGTCAGCAAG 275
Db      233 TCGGACGCCAAGCTACATGAGCATGTGCTCCAGTCCCAAGAGATGATGAAGAC 292
Qy      276 CGAAGATGACGGCTTAACTGAGCAAGCACTCAGGCTCTCAGGCGACATCTGCCACAG 335
Db      293 AGGAGAGGCCAAGCTGGCTG----- 313
Qy      336 CTGCTGCTGTTGAGCCCCCGCTCTCTGATCATGAGCAAGCTCATGCGAGTTGAGA 395
Db      314 CCTCTCTATGACAGCATGCTCAATGCGACGATGCTGTCAATGAAACAGTTGTTGCTG 373
```

```
Qy      396 GAACAATGATTGTCGCGATCAATCTTCAGCTATCAACAAGACTGGT 455
Db      374 GCACACAGTGGTGGCTGCTGCTCTCTCAATGAGTTGCAACAATGAGCT 433
Qy      456 CAGTCCCTCAGCAGTGAAGTCTCTATGCAAGTCTCTAGACC 515
Db      434 ATGTTGCTGAGACTAGGAGGGGGGGGGGTATCTCCAAAGGAAAGT 493
Qy      516 GATGCGGATCTTGGCAGAAAGCGTTGGATTGACAGATTAAAGTGTG 575
Db      494 GGTGACGACTTGGCAGAAAGGTTGGACTTACAGATTCAACTGCG 553
Qy      576 TCTACTGGGCTTTACATGGTACTCGGACAAACACACTTGCATATGA 635
Db      554 TGTACTGTGACTGACCGCTACTCCGACACGACGACTGCAAGTTGCA 613
Qy      636 CAGGCGAGGAAGCGATTGGCGAAAGCTAATCTCTTGTGTCGGAGAA 695
Db      614 CTGCTAGGGATGCCATTGGCAGAGGCTAATCCAGTGTGAAGCGGACAA 673
Qy      696 TTT 698
Db      674 TCT 676
```

```
RESULT 11
US-10-425-114-25774
; Sequence 25774, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tadaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improv
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25774
; LENGTH: 932
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3959-015-A1_FLI
US-10-425-114-25774
```

```
Query Match      12.8%; Score 102.6; DB 7; Length 9
Best Local Similarity 52.9%; Pred. No. 7.4e-24;
Matches 337; Conservative 0; Mismatches 264; Indels 4;

Qy      82 TTTTCATACGCTTCTGATTAAGAAAGATACACCTTATTAAGTCTGTA 141
Db      136 TGTGACATCACTATTCTGTGATATTGATCACTTGTGTTGATTAATTCAGTT 195
Qy      142 CGAGGTGTGCTCAGAGACGACCTCGAGGCCCTTAGGGTCCAGTT 201
Db      196 GAGC-----ACAAGAGAGCTGGGTGTCAGGCCCCCGAGGAGCAATC 249
Qy      202 CCTTTGCGGCTTTCTTGGCAGCCAAAGCTACATGAGGTTGTGCTGAG 261
Db      250 TAACTGTGCTTTCTTGGCAGGCGGCGACCATGAACATGTGCTCAAG 309
Qy      262 GACAGTCATGCAAGGAGATGACGCGCTTTAGCTGAGCAAGCACTGAC 321
Db      310 GATGATTAAGAAAGCAGATACAGGCCAAGCTGGCTG----- 355
Qy      322 GACATTCGCAAGCTGCTGCTTTCAAGCCCCCGCTCTCTGATGAG 381
```



```

Db      356 GACAGCATGCTGAAAGCGACGACGCCGT-----CATGAGCCGGTTGTTG 401
Qy      382 ATCGAGGTTGAGAGAAACAATGATTGTGCCGATCAATCTTCCAGTTAACAAGACCT 441
Db      402 CTGGCAGCAACACGGTAGTAGCTGTGTGCCCAGTTAGTTGCAACAAAGAAC--GTGCA 459
Qy      442 GGTAAACCCCGCTGACGCTGCCCTCAGGCAAGTAAGTCTCTATGCGAGCTCCCTCTAG 501
Db      460 GCAAGCCCGCTGATGTTGTCGACGACCAACGAGAGGGGTGGCGGATCTTCAAAAGGGGAA 519
Qy      502 ACCCGAGCCCAATCGATGCGGATCTTTCAGAGAACCTGTGTGATTGACAGATTAAAGT 561
Db      520 GGTAGGGCCGAACCGGTGACGGGCTGCAAGAAAGGTTGACTTACGGATTCAACTG 579
Qy      562 TGGCTGTGCAACCTTACTGCGCTTTTACATGCGTACTGGAACAACACATTGACATA 621
Db      580 CCGGTGTGGAACTGTACTGCGCACTCCACCGCTACTCCGACAGACAGATGCAAGTT 639
Qy      622 TGACTACAAAGCCGACGAGGAGGAGGATGCGAAAGCTTATCCCTTGTGTCGCGCA 681
Db      640 CGACTACCGGACTGCTGCCAGGAGCGCAATGCGAAAGCTAAATCCGGTGTGAAGGACA 699
Qy      682 GAAGTTGTCAAGTTTGTATGAGCATCCGTTAAGCTT 718
Db      700 CAAGCTCGACAAAGATCTAGGGGGGTTCCCTACGGTT 736

```

```

RESULT 12
US-10-425-115-71464
; Sequence 71464, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 71464
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_16517C.1
US-10-425-115-71464

```

```

Query Match      12.8%; Score 102.6; DB 8; Length 1114;
Best Local Similarity 52.9%; Pred. No. 8.1e-24;
Matches 337; Conservative 0; Mismatches 264; Indels 36; Gaps 4;

Qy      82 TTTCTCATCGCTTCTGATTAGAAAGATACACCTTATTACTCTTTAAAGATGCTCAC 141
Db      136 TGTCACTACTATATTTGGATTGATCACTGTTTGGTATTATTTAGTTAAGAAAGCAAT 195
Qy      142 CGAGCGTGTGTCAGAGACGACCTCGCAGGCCCCCTGAGGGTCCAGTTATGTGCAAGA 201
Db      196 GGAGC-----ACAAGAGAGGCTGGGTGCAAGGCCCCCAAGGAGCAACCTCTGTGATCAA 249
Qy      202 CTTTTCGGCTTTCTTGGGAGCCAAAGCTACATGAGGGTTGTGCTGAAAGTCTACCGAGA 261
Db      250 TAACTGTGGCTTTCTTGGGAGCCGCGGACACATGAACATGTCTCCAAGTGCACAAAGAA 309
Qy      262 GACAGTCATGCAAGGAGAGATGACGGCTTTAGTACAGCAAGCATCAGGCTCAGGCG 321
Db      310 GATGATTAACGAAGCAGAGATCAAGCCCAAGCTGAGTGC-----CTCTCTATATC 355
Qy      322 GACATCTGCAAGCTGCTGTCTTCAAGCCCCCGCTCTGTATGATGAGAACCAAGCTCAC 381
Db      356 GACAGCATGCTGTAACGCGACGACGCCGT-----CATGAGCCGGTTGTTG 401

```

```

Qy      382 ATCGAGGTTGAGAGAAACAATGATTGTGCCGATCAATCTTCCAGTTAACAAGACCT 441
Db      402 CTGGCAGCAACACGGTAGTAGCTGTGTGCCCAGTTAGTTGCAACAAAGAAC--GTGCA 459
Qy      442 GGTAAACCCCGCTGACGCTGCCCTCAGGCAAGTAAGTCTCTATGCGAGCTCCCTCTAG 501
Db      460 GCAAGCCCGCTGATGTTGTCGACGACCAACGAGAGGGGTGGCGGATCTTCAAAAGGGGAA 519
Qy      502 ACCCGAGCCCAATCGATGCGGATCTTTCAGAGAACCTGTGTGATTGACAGATTAAAGT 561
Db      520 GGTAGGGCCGAACCGGTGACGGGCTGCAAGAAAGGTTGACTTACGGATTCAACTG 579
Qy      562 TGGCTGTGCAACCTTACTGCGCTTTTACATGCGTACTGGAACAACACATTGACATA 621
Db      580 CCGGTGTGGAACTGTACTGCGCACTCCACCGCTACTCCGACAGACAGATGCAAGTT 639
Qy      622 TGACTACAAAGCCGACGAGGAGGAGGATGCGAAAGCTTATCCCTTGTGTCGCGCA 681
Db      640 CGACTACCGGACTGCTGCCAGGAGCGCAATGCGAAAGCTAAATCCGGTGTGAAGGACA 699
Qy      682 GAAGTTGTCAAGTTTGTATGAGCATCCGTTAAGCTT 718
Db      700 CAAGCTCGACAAAGATCTAGGGGGGTTCCCTACGGTT 736

```

```

RESULT 13
US-10-424-599-61443
; Sequence 61443, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules and Other Mol-
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro-
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 61443
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26497C.1
US-10-424-599-61443

```

```

Query Match      12.7%; Score 102.4; DB 7; Length 1191;
Best Local Similarity 70.8%; Pred. No. 9.8e-24;
Matches 136; Conservative 0; Mismatches 56; Indels 0;

Qy      509 CCCATGATGCGATCTTTCAGAGAGAGGCTGTGATTGACAGATTTTAA 568
Db      617 CCAAAACGCTCAGCAGCTGCAACAGGAGTGTGTTGACAGGCTTTAA 676
Qy      569 GGCACCTCTACTGCGCTTTTACATCGGTAATCGGACAAACACACTTGC 628
Db      677 GGTGACCTTTTGTGCTGTAATCGTACTACAGCAAGCAATATTTGCC 736
Qy      629 AAAGCCGACGAGGAGAAAGCAATTTGCGAAAGCTAATCTTGTGTGAG 688
Db      737 CGCACTGCTGCTCGGAGATGCAATAGCTTAAACAAACCCAGTTGCAAGG 796
Qy      689 GTCAAGTTTGA 700
Db      797 GATTAAGATCTAA 808

```

```

RESULT 14
US-10-425-114-27574
; Sequence 27574, Application US/10425114

```

Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jinding
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 27574
LENGTH: 1032
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB4732-053-All_FLI
US-10-425-114-27574
Query Match 12.7%; Score 102.2; DB 7; Length 1032;
Best Local Similarity 51.2%; Pred. No. 1.1e-23;
Matches 288; Conservative 0; Mismatches 248; Indels 27; Gaps 1;
QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTTATGTGCAAGAACCTTGGCGCTTCT 215
DB 135 AGGAGGCTGGGTGCGAGGCCCCGAGGAGCCCATCTCTGCAATCAATATGTGGCTTCT 194
QY 216 TCGGACGCAAGCTACATGAGGTTGTGCTCGAATGTGCTACCGAGAGACATGATCAAG 275
DB 195 TCGGACGCGCGCGACATGAAATGTGCTCAATGTGCGACAGAGATGATGAAGAGC 254
QY 276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGCGCATGTGCGACAG 335
DB 285 AGGA-----TCAGGCCAAGCTGGCTCGCTCTCTATTTG 287
QY 336 CTGCTGCTGTTAGGCCCCCGCTCTGTATCATAGAACCAAGCTCATGCGAGTTGAGA 395
DB 288 ACAGCATGCTGAACGCGACGAGCGCGTCATGAGACCGGTGGTGTGCGAGCAACACGG 347
QY 396 GAACATGATTTGTGCGCGCATCATCTTCAGCTATCAACAAGACCTGGTTACCCCGCTG 455
DB 348 TAGTAGCTGTGGCCCAAGTCGAGTTGCAAAACAATGAACGTGACAGACCCGCTGATTTG 407
QY 456 CAGTGGCCCTCAGGCAAGTGAAGTCTTATGCAAGCTCTCTTACGCCAGCCCAATC 515
DB 408 CCGGACCGACGCGAGGGGGTGGCGCGCGATCTCCAAAGGGGGAAGGTAGGGCCGAAC 467
QY 516 GATGGGATCTTTGCAAGAACGCTGTGATTTGACAGATTGAAGTGTGGCTGGCAACC 575
DB 468 GGTGACGCGCTGCAAGAAAGGTTGGCTCAGGAGTTCAACTGCGGTGTGGAACT 527
QY 576 TCTACTGGGCTTTACATCGGTACTCGACAACAACACTTGCATATGACTCAAAAGCG 635
DB 528 TGTATGCGCGCTTCAACGCTACTCGACAACAAGACTGCAAGTTGCACTACCGGACTG 587
QY 636 CAGGACAGAAAGCATTTGCGAAAGCTAATCTTGTGCTGGCCGAGAAAGTTGCAAGT 695
DB 588 CTGCCAGGAGCGCATTTGCGCAAGGCTAATCCGTTGTGAGAGCAGCAAGCTCGACAAGA 647
QY 696 TTTGATGAGCATCCGTTAAGCTT 718
DB 648 TCTAGGGGGGGTTCCTTACGGTT 670

RESULT 15
US-10-425-114-22047

Sequence 22047, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jinding

APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 22047
LENGTH: 736
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3356-023-F11_FLI
US-10-425-114-22047

Query Match 12.7%; Score 101.8; DB 7; Length 72
Best Local Similarity 53.5%; Pred. No. 1.2e-23;
Matches 301; Conservative 0; Mismatches 232; Indels

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTTATGTGCAAGAACCTT 215
DB 114 AGGAGGCTGGGTGCGAGGCCCCGAGGAGCCCATCTCTGCAATTA 173
QY 216 TCGGACGCAAGCTACATGAGGTTGTGCTCGAATGTGCTACCGAGAGC 275
DB 174 TCGGACGCGCGCGACATGAATGTGCTCAATGTGCGACAGAGAT 233
QY 276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGCGAC 335
DB 234 AGGATCAGGCCAAGCTGCTGCT-----CTCTCTATTCAGC 279
QY 336 CTGCTGCTGTTAGGCCCCCGCTCTGTATCATGAGACCAAGCTCATG 395
DB 280 CCGGACGCGCGCGT-----CATGAGCGCGTTGTTGCTG 325
QY 396 GAACATGATTTGTGCGCGCATCATCTTCAGCTATCAACAAGACCTG 455
DB 326 GTAGTAGCTGTGGCCCAAGTTGAGTTGCAACAATGAAC--GTGACAGA 383
QY 456 CAGTGGCCCTCAGGCAAGTGAAGTCTTATGCAAGCTCTCTTACACC 515
DB 384 TTGCGGACCGACGAGGGGGTGGCGCGATCTCCAAAGGGGGAAGT 443
QY 516 GATGCGATCTTTGCAAGAACGCTGTGATTTGACAGATTTAAGTGTG 575
DB 444 GGTGACGCGCTGCAAGAAAGGTTGGACTTACGGGATTCACATGCGCG 503
QY 576 TCTACTGGGCTTTACATCGGTACTCGACAACAACACTTGCATATTA 635
DB 504 TGTATGCGGACTCCACGCTACTCCGACAAGACGACTGCAAGTTGCA 563
QY 636 CAGGACAGAAAGCATTTGCGAAAGCTAATCTTGTGCTGGCCGAGAA 695
DB 564 CTGCCAGGAGCGCATTTGCCAAGGCTAATCCGTTGTGTAAGGACAGACA 623
QY 696 TTTGATGAGCATCCGTTAAGCTT 718
DB 624 TCTAGGGGGGGTTCCTTACGGTT 646

Search completed: December 8, 2005, 10:29:11
Job time : 803 secs


```

; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 47138
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Bovine 19866881619274
US-10-750-185-47138

Query Match      4.8%; Score 38.2; DB 6; Length 1067;
Best Local Similarity 54.7%; Pred. No. 0.0067;
Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1
567 GTGCAACCTTACTGCGCTTACATCGGTACTCGACAACACTTGCACATATGACT 626
    |||||
574 GTGGAATGTTTACTGTAGTGTACACATTACTCAGATGTACACAACAGCTCTTACAATT 515
    |||||
627 ACAAGCCGACGACGACGACGAGGATTCGGAAGCTAATCTTTCGCGCGCGAAGG 686
    |||||
514 AAGACTGATGCTGCTGTGAAAATCAGAAAATCAGAAATCAGTACTTGTGTGAAAAGA 455
    |||||
687 TTGTCAAGTTTGTATGATGAC 705
    |||||
DB 454 TCCAGAGATTGTGAAGCTGC 436

RESULT 3
US-11-153-880-3/C
; Sequence 3, Application US/1153880
; Publication No. US20050256050A1
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/153,880
; FILING DATE: 16-Jun-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/219,442
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/999,811
; FILING DATE:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWICZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488.1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1526 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 71..142
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 143..1120
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 71..1120
US-11-153-880-3

Query Match      4.1%; Score 32.8; DB 7; Length 15;
Best Local Similarity 61.9%; Pred. No. 0.6;
Matches 52; Conservative 0; Mismatches 32; Indels 0;

QY
23 AGTGTGTGTCTTACTAGTGTGTTGCAAGCTGAAGCTGAGGAGAGA 82
    |||||
DB 484 AGTGTATTTGGCAAACTGATTTTACTGTTTGGGCTTGAGAGAGA 425
    |||||
QY 83 TTTCATACGCTTCGATTAGGAA 106
    |||||
DB 424 TTCAATTAAGCTTTCCTGAGGTA 401

RESULT 4
US-11-153-880-1/C
; Sequence 1, Application US/1153880
; Publication No. US20050256050A1
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/153,880
; FILING DATE: 16-Jun-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/219,442
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/999,811
; FILING DATE:

```

FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1674 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 12..80
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 81..1268
FEATURE:
NAME/KEY: CDS
LOCATION: 12..1268
US-11-153-880-1

Query Match 4.1%; Score 32.8; DB 7; Length 1674;
Best Local Similarity 61.9%; Pred. No. 0.63; Indels 0; Gaps 0;

Matches 52; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 23 AGTGTGTGTCATTAGTGTGTCGAAGCTTGAGCGAGATTGCAAGATT 82
DB 632 AGTGTGATGGCAAACTGATTGTTACTGTTGGGGCTTGAGAGAGGCACTGTAT 573

QY 83 TTCTCATACGCTTCGATTAGGA 106
DB 572 TTCAATTAACGCTTGTGAGGTA 549

RESULT 5

US-11-064-769-1/c
Sequence 1, Application US/11064769
GENERAL INFORMATION:
APPLICANT: Alltalo, Karl
APPLICANT: Hilunen, Mikko O
APPLICANT: Jeltsch, Markku M
APPLICANT: Achen, Marc G
TITLE OF INVENTION: Use of VEGF-C or VEGF-D Gene or Protein to Prevent Restenosis
FILE REFERENCE: 28967/35601A
CURRENT APPLICATION NUMBER: US/11/064,769
CURRENT FILING DATE: 2005-02-24
PRIOR APPLICATION NUMBER: US/09/427,657
PRIOR FILING DATE: 1999-10-26
PRIOR APPLICATION NUMBER: US 60/105,587
PRIOR FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1997
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (352)..(1608)
US-11-064-769-1

Query Match 4.1%; Score 32.8; DB 7; Length 1997;
Best Local Similarity 61.9%; Pred. No. 0.7;

Matches 52; Conservative 0; Mismatches 32; Indels 0;
QY 23 AGTGTGTGTCATTAGTGTGTCGAAGCTTGAGCGAGATTGCAAGATT 82
DB 972 AGTGTGATGGCAAACTGATTGTTACTGTTGGGGCTTGAGAGAGATT 913
QY 83 TTCTCATACGCTTCGATTAGGA 106
DB 912 TTCAATTAACGCTTGTGAGGTA 889

RESULT 6

US-11-064-774A-21/c
Sequence 21, Application US/11064774A
GENERAL INFORMATION:
APPLICANT: Alltalo et al.
TITLE OF INVENTION: MATERIALS AND METHODS INVOLVING HYBRID V
FILE REFERENCE: 28967/35977B2
CURRENT APPLICATION NUMBER: US/11/064,774A
CURRENT FILING DATE: 2005-02-24
PRIOR APPLICATION NUMBER: 09/795,006
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: US 60/205,331
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 60/185,205
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 1212
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 21
LENGTH: 1997
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (352)..(1608)
US-11-064-774A-21

Query Match 4.1%; Score 32.8; DB 7; Length 1997;
Best Local Similarity 61.9%; Pred. No. 0.7;
Matches 52; Conservative 0; Mismatches 32; Indels 0;

QY 23 AGTGTGTGTCATTAGTGTGTCGAAGCTTGAGCGAGATTGCAAGATT 82
DB 972 AGTGTGATGGCAAACTGATTGTTACTGTTGGGGCTTGAGAGAGATT 913
QY 83 TTCTCATACGCTTCGATTAGGA 106
DB 912 TTCAATTAACGCTTGTGAGGTA 889

RESULT 7

US-11-064-769-17/c
Sequence 17, Application US/11064769
GENERAL INFORMATION:
APPLICANT: Alltalo, Karl
APPLICANT: Hilunen, Mikko O
APPLICANT: Jeltsch, Markku M
APPLICANT: Achen, Marc G
TITLE OF INVENTION: Use of VEGF-C or VEGF-D Gene or Protein
FILE REFERENCE: 28967/35601A
CURRENT APPLICATION NUMBER: US/11/064,769
CURRENT FILING DATE: 2005-02-24
PRIOR APPLICATION NUMBER: US/09/427,657
PRIOR FILING DATE: 1999-10-26
PRIOR APPLICATION NUMBER: US 60/105,587
PRIOR FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 17

LENGTH: 2679
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: chimeric
OTHER INFORMATION: sequence in which CMV promoter sequence is ligated
OTHER INFORMATION: to Homo sapien VEGF-C sequence
US-11-064-769-17

Query Match 4.1%; Score 32.8; DB 7; Length 2679;
Best Local Similarity 61.9%; Pred. No. 0.83;
Matches 52; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 23 AGTGTGTGTCTATTAGTCTTGTGCAAGCTTGAAGCGGAGATTGGCAGATT 82
DB 1654 AGTGTGATTGGCAAAACGATGTTACTGTGGGCGCTTGAAGAAGGCACGTGAT 1595
QY 83 TTCTCATACGCTTCTGTGATTAGGAA 106
DB 1594 TTCAATTAACGTCTTGTGAGGTA 1571

RESULT 8
US-11-121-086-75/c
Sequence 75, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 75
LENGTH: 161874
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-75

Query Match 4.0%; Score 32; DB 7; Length 161874;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 80; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 525 CTTGACGAAAGCGTGTGTGATGACAGATTAAAGTGTGCGCAACTCTACTGCG 584
DB 131884 CTGACATGAAGCGTGGGATTTGATTTAGTAGTTTCAGGGTGAGCACAGAACTTG 131825
QY 585 CTTTACATCGTACTCGGCAACAACACTTGCACATATGACTACAAAGCCGAGGCGAG 644
DB 131824 CTTATGATCAGTCAGCTAACACATACACATAGAAATACAGCAAGTTGAGGTAGAG 131765
QY 645 AAGCGATTGCCAAAGCTATCTCTTGTGTGCGCGAAGA 684
DB 131764 AGACAAAGTGAAGAAAGCTTATGCTATCATGTCACGAGTGAGAA 131725

RESULT 9
US-11-121-086-20/c
Sequence 20, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04

NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 20
LENGTH: 162289
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-20

Query Match 3.9%; Score 31.6; DB 7; Length 16.
Best Local Similarity 49.1%; Pred. No. 23;
Matches 112; Conservative 0; Mismatches 114; Indels 1;

QY 23 AGTGTGTGTCTATTAGTGTGTTGCAAGCTTGAAGCGCTTGAAGGAGA 82
DB 91649 AGAGTACTGTCTTGTCTGACTGACCTTGTGTGAATATCTAATAAGC 91590
QY 83 TTCTCATACGCTTCTGTATTAGAAAGATACACCTTATTAATCTGTAA 142
DB 91589 CTCTGCCATGGCTCTGGCTAAGTGGAGGCCCGTGTGGTCTGTGAC 91530
QY 143 GAGCGTGTCT--TCAGAGACGACCTGCGAGGCCCTCGAAGGTTCCACT 200
DB 91529 CAGAGCTGTCTCTTGTGACTTGAAGGCTGAGGCGCTGCCAGCGCAC 91470
QY 201 ACCTTGGCGCTTCTTGTGCGACCAAGCTTACATAGGAGTTGTCTGA ;
DB 91469 GCCTCAGTCTCTCATCTGTGACAAATAATACAGGCGCTGTGCCCCA ;

RESULT 10
US-11-121-086-35
Sequence 35, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 35
LENGTH: 170995
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-35

Query Match 3.9%; Score 31.2; DB 7; Length 17.
Best Local Similarity 50.7%; Pred. No. 33;
Matches 75; Conservative 0; Mismatches 73; Indels 0;

QY 152 TCTGAGAGACGACCTGCGAGGCCCTGAGGATCCAGTTATATGCAAGA 211
DB 32928 TCTTGCACCCCAAGTCCCAACCAAGCAAGGGGCGCTCACTCGACA 32987
QY 212 TTCTTGGCAGCCAAAGCTACATGAGGTTGTCTGAAAGCTACCGAG 271
DB 32988 TTCCGGGGGGCGCCGAAGGCCCAAGGGGTCTGGGGGGAATTGCAATTGC 33047
QY 272 CAAGGAAGATGACGGCTTTAGCTGAGC 239
DB 33048 GGAGCGAGCTGAAGTGTCTCGCGGAGC 33075

RESULT 11
US-10-750-185-46945/c
Sequence 46945, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.

APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 46945
LENGTH: 2070
TYPE: DNA
ORGANISM: Bovine 19866880571876
US-10-750-185-46945

Query Match 3.9%; Score 31; DB 6; Length 2070;
Best Local Similarity 62.0%; Pred. No. 3;
Matches 49; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 350 CCCCCGCTCTGTACATGAGCCAGCTCAGATGCGAGTTGAGAGACATGATGTG 409
DB 2050 CCCCCATCTCCATCATTAAGATTAACCTGACATCCACCCTGACAGACAGTATTG 1991
QY 410 CCGCATCATCTTCCAGCT 428
DB 1990 AGACATTAGTTGCCATCT 1972

RESULT 12
US-11-121-086-58
Sequence 58, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO: 58
LENGTH: 180654
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-58

Query Match 3.9%; Score 31; DB 7; Length 180654;
Best Local Similarity 43.7%; Pred. No. 40;
Matches 79; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 587 TTACATCGGACTCCGACAAACACACTTGACATATGACTACAAAGCCGACGACGAA 646
DB 29799 TTTCCAAACAAATGGGGCTAACTCTCCCTCAATGACTTGAAGAGCCGAGGAAATTA 29858
QY 647 GCGATTGGAAAGCTAATCTCTTGTCTGCGCCGAGAAAGTTGTCAAGTTTGTATGAGA 706
DB 29859 TGGATTACTATATTAACCTCTTACCTGTCTATATAAGGAGCAATTAAGAACTGGA 29918
QY 707 TCCGTTAAGCTTTGCGCCGACAGATTTAGGCTTCATACA 745
DB 29919 TCATTAAAGCTTGATCTTCATTAATTTGGTACATTA 29957

RESULT 13
US-10-750-185-54345/c

Sequence 54345, Application US/10750185
Publication No. US2005026063A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 54345
LENGTH: 1433
TYPE: DNA
ORGANISM: Bovine 19866880647140
US-10-750-185-54345

Query Match 3.8%; Score 30.2; DB 6; Length 14;
Best Local Similarity 55.1%; Pred. No. 4.6;
Matches 59; Conservative 0; Mismatches 48; Indels 0;

QY 657 AAGTAATCCTCTTGTGCGCCGAGAGGTTGTCAAGTTTGTATGAGC 716
DB 828 AAGTAATCATGAAGTTGATGCTGATTAAGACAGTTTGGTTTGAAGACT 769
QY 717 TTTTCTGCCGAGATTAGGCTTCATACATTGAGTACTTACATCT 716
DB 768 TTTCTCGAAGTAATTTGGTCAATAAAGTGAATATGACATTTGCT 769

RESULT 14
US-10-750-185-33374
Sequence 33374, Application US/10750185
Publication No. US2005026063A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 33374
LENGTH: 1523
TYPE: DNA
ORGANISM: Bovine 19866881033634
US-10-750-185-33374

Query Match 3.7%; Score 30; DB 6; Length 1523;
Best Local Similarity 54.5%; Pred. No. 5.5;
Matches 60; Conservative 0; Mismatches 50; Indels 0;

QY 122 AGTCTGTTAAGATGCGACCGAGCGGTGTCTCAGAGACGACCTGC 181
DB 889 AGTATATGTGAGATGCTCTGGAGAGGAGCAATGAGAGGCTGCTG 948
QY 182 GGTCCAGTTATGTGCAAGACCTTTGGGGCTTCTTCCGACGCAAGCTA 948

Db 949 ATTCTTCTTTTGCATCATCATTAATGCTTGCAGATGGTCAATGCTCC 998

RESULT 15

US-10-750-185-56847
; Sequence 56847, Application US/10750185
; Publication No. US20050260603A1
GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KER, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INVERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56847
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Bovine 19866880892313
US-10-750-185-56847

Query Match 3.7%; Score 29.8; DB 6; Length 837;

Best Local Similarity 54.0%; Pred. No. 4.6;
Matches 61; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 502 ACCGAGCCCAATCGATCGGATCTTGACAGAGCGTGTGATTCACAGATTAAAGTG 561
Db 450 ACGGGGGCCCAATCTTCTGACTGTGTGCGAGGAAATGTCTTCTTGGAACGATTGACAGGG 509
OY 562 TCGCTGTGGCAACCTCTACTGCGCTTTACATCGGTACTCGACAAACACACTT 614
Db 510 CCACTGGGGAAGCTAAGCAGCCCTTGATTCTTTCTTGCGTGGCAACTTT 562

Search completed: December 8, 2005, 10:33:29
Job time : 250 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 12:54:20 ; Search time 134 Seconds
(without alignment)
616.442 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSGEETTSQARPECPV.....GGEALAKANPLVAKVYKF 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: A_Geneseq_21:*
- 2: geneseqp1980s:*
- 3: geneseqp1990s:*
- 4: geneseqp2000s:*
- 5: geneseqp2001s:*
- 6: geneseqp2002s:*
- 7: geneseqp2003as:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description.
1	984	100.0	188	5	ABG93809	Abg93809 Transcrip
2	984	100.0	188	5	AAU71800	AAU71800 Zinc fing
3	416.5	42.3	173	7	ABM90068	Abm90068 Rice abio
4	416.5	42.3	173	7	ABM87792	Abm87792 Rice abio
5	412.5	41.9	179	8	ADY97065	Ady97065 Plant ful
6	412.5	41.9	179	8	ADY09608	Ady09608 Plant ful
7	408.5	41.5	169	7	ABM73978	Abm73978 DNA clone
8	402.5	40.9	173	8	ADY75512	Ady75512 Plant ful
9	402.5	40.9	173	8	ADY94337	Ady94337 Plant ful
10	401.5	40.8	173	8	ADY95012	Ady95012 Plant ful
11	386	39.2	160	7	ABM90052	Abm90052 Rice abio
12	386	39.2	169	7	ABM87359	Abm87359 Rice abio
13	379	38.5	173	3	AAQ32536	AAQ32536 Arabidops
14	379	38.5	173	3	AAQ49664	AAQ49664 Arabidops
15	379	38.5	173	3	AAQ49652	AAQ49652 Arabidops
16	378	38.4	173	3	AAQ23282	AAQ23282 Arabidops
17	374	38.0	173	3	AAQ54935	AAQ54935 Arabidops
18	366	37.2	213	2	AA422772	AA422772 Rat neuro
19	365	37.1	213	4	ABU52627	ABU52627 Human bra
20	365	37.1	229	5	ABP41626	ABP41626 Human ova
21	365	37.0	213	3	AAQ56766	AAQ56766 Human pro
22	364	37.0	213	4	AAQ64505	AAQ64505 Gene 24 h
23	364	37.0	213	4	AAQ98228	AAQ98228 Human zin
24	364	37.0	213	7	ABG75067	ABG75067 Human zin

25	364	37.0	213	8	AD124504	Ad1:
26	364	37.0	213	8	ADN04776	Adn:
27	364	37.0	213	8	ADP24800	Adp:
28	364	37.0	213	9	ADP07989	Adx:
29	364	37.0	213	9	AEA98119	Aea:
30	362.5	36.8	170	3	AAQ14725	Aag:
31	361.5	36.7	175	3	AAQ43065	Aag:
32	361.5	36.7	175	3	AAQ17575	Aag:
33	359.5	36.5	170	3	AAQ17628	Aag:
34	359	36.5	190	7	ABM86938	Abm:
35	357	36.3	224	8	ADY23911	Ady:
36	356	36.2	161	3	AAQ43290	Aag:
37	356	36.2	161	3	AAQ09127	Aag:
38	355	36.1	233	8	ADP22660	Adp:
39	352	35.8	223	7	ADP59857	Adp:
40	351.5	35.7	170	3	AAQ16569	Aag:
41	351	35.7	205	8	ADP22662	Adp:
42	346.5	35.2	163	3	AAQ30169	Aag:
43	346.5	35.2	163	3	AAQ49941	Aag:
44	346.5	35.2	163	3	AAQ49016	Aag:
45	346.5	35.2	199	3	AAQ49940	Aag:

ALIGNMENTS

RESULT 1	ABG93809	standard; protein; 188 AA.
ID	ABG93809	
XX	ABG93809;	
AC		
DT	28-NOV-2002 (first entry)	
XX		
DE	Transcription factor stress-related protein (TFSRP) #4.	
XX		
KW	Transcription factor stress-related protein; TFSRP; MYB-1	
KW	AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;	
KW	ZF-4 protein; ZF-5 protein; CAAT-Box binding factor prote-	
KW	Sigma factor like protein; SFL-1; transgenic; plant; drou	
KW	environmental stress; Physcomitrella patens; high salinity	
KW	ciliate; fungus.	
XX		
OS	Physcomitrella patens.	
XX		
PN	US2002102695-A1.	
XX		
PD	01-AUG-2002.	
XX		
PF	06-APR-2001; 2001US-00828303.	
XX		
XX	07-APR-2000; 2000US-0196001P.	
PA	(SILV/) SILVA O D C E.	
PA	(BOHN/) BOHNERT H J.	
PA	(THIE/) THIELEN N V.	
PA	(CHEN/) CHEN R.	
PI	Silva ODE, Bohnert HJ, Thiele N, Chen R;	
XX		
DR	WPI; 2002-690614/74.	
DR	N-PSDB; ABS70622.	
PT	Novel transcription factor stress-related protein such as	
PT	protein, zinc-finger protein, or CAAT-Box binding factor f	
PT	useful for modifying stress tolerance of plant.	
XX		
PS	Claim 2; Fig 3D; 104dp; English.	
XX		
CC	The invention relates to an isolated transcription factor	
CC	protein (TFSRP) (I) selected from AP2 similar-2 protein (A	
CC	finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein	
CC	MYB-1 protein, CAAT-Box binding factor protein-3 (CABP-3),	
CC		

CC like protein (SFL-1) and their orthologues. Also described is: (1) an
CC isolated TFSRP coding nucleic acid (II) coding for (I); (2) a transgenic
CC plant cell (III) transformed by (II), where the expression of (II) in the
CC plant cell results in increased tolerance to an environmental stress as
CC compared to a wild-type variety of the plant cell; (3) a transgenic plant
CC (IV) comprising (III); (4) a seed (V) produced by (IV), where the seed is
CC true breeding for an increased tolerance to environmental stress as
CC compared to a wild-type variety of the plant cell; (5) an agricultural
CC product (VI) produced by (IV) or (V); and (6) an isolated recombinant
CC expression vector (VII) comprising (II), where expression of (VII) in a
CC host cell results in increased tolerance to environmental stress as
CC compared to a wild-type variety of the host cell. (II) is useful for
CC modifying (increasing or decreasing) stress tolerance of a plant
CC (transgenic or not transgenic), by modifying the expression of TFSRP in
CC the plant. The plant is transformed with a promoter that directs
CC expression of the TFSRP. The promoter is tissue specific and is
CC developmentally regulated. TFSRP expression is modified by administration
CC of an antisense molecule that inhibits expression of TFSRP. (VII) is
CC useful for producing a transgenic plant containing (II), where expression
CC of the nucleic acid in the plant results in increased tolerance to
CC environmental stress as compared to a wild-type variety of the plant, by
CC transforming a plant cell with (VII) comprising the nucleic acid,
CC generating from the plant cell a transgenic plant with an increased
CC tolerance to environmental stress as compared to a wild-type variety of
CC the plant. (I) is useful for conferring stress tolerance such as drought,
CC cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful
CC for identifying Physcomitrella patens and related organisms, for mapping
CC of genomes of organisms related to P.patens, for identifying and
CC localizing P.patens sequences of interest, for evolutionary studies, for
CC determining TFSRP regions required for function, for modulating TFSRP
CC activity, for modulating metabolism of one or more cell functions, for
CC modulating transmembrane transport of one or more compounds, and for
CC modulating stress resistance. (I) is useful for reducing stress tolerance
CC response plants or more particularly, in the transcription of a protein
CC involved in a stress tolerance response in a P.patens plant. (II) is
CC useful for transforming plants and thus inducing tolerance to stresses
CC such as drought, high salinity and cold, for identifying the presence of
CC P.patens or a related organism in a mixed population of microorganism,
CC serve as marker for specific regions of the genome for mapping the genome
CC and for the functional studies of P.patens proteins. (II) is useful as
CC reference points for mapping the moss genomes, or of genomes of related
CC organisms, for evolutionary and protein structural studies, for
CC generating knockout mutation in the genomes of various organisms such as
CC bacteria, mammalian cells, yeast cells and plant cells, useful for
CC evaluating their ability or capacity to tolerate various stresses
CC conditions and the effect on the phenotype and/or genotype of the
CC mutation. (II) is useful as marker for the construction of a genomic map
CC in related mosses. (I) or (II) is useful for generating algae, ciliates,
CC plants, fungi or other microorganisms expressing mutated TFSRP nucleic
CC acid and protein molecules for improving stress tolerance. ABG93806-
CC ABG93813 represent P. patens TFSRP amino acid sequences of the invention
XX
XX
SQ Sequence 188 AA;

Query Match 100.0%; Score 984; DB 5; Length 188; 1
Best Local Similarity 100.0%; Pred. No. 6.9e-92; Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATERVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLGSKCYRETYWQAKMTALAEQATCA 60
DB 1 MATERVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLGSKCYRETYWQAKMTALAEQATCA 60
QY 61 AQAATATAAANVPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
DB 61 AQAATATAAANVPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
QY 121 PSRPPNRCGSKRKVGLTGFKRCGNLYCALHRYSDKHTCYVDYKAAQGEAIAANPLV 180
DB 121 PSRPPNRCGSKRKVGLTGFKRCGNLYCALHRYSDKHTCYVDYKAAQGEAIAANPLV 180
QY 181 VAEKVVKF 188
DB 181 VAEKVVKF 188

RESULT 2
AAU71800
ID AAU71800 standard; protein, 188 AA.
XX
AC AAU71800;
XX
DT 26-FEB-2002 (first entry)
XX
DE Zinc finger-4 (ZF-4).
XX
KW Transcription factor stress-related protein; AP2 similar-2
KW AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;
KW MYB-1 MYB-1 protein, CAAAT-box binding factor protein-3; CA
KW Sigma factor like protein, SFL-1; environmental stress; mo
XX
OS Physcomitrella patens.
XX
PN WO200177311-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US011393.
XX
PR 07-APR-2000; 2000US-0196001P.
XX
PA (BAD I) BASF PLANT SCI GMBH.
XX
PI Da Costa Silva O, Bohnert HJ, Van Thielten N, Chen R;
XX
DR MPI, 2002-049146/06.
DR N-FSDB; AAS95832.
XX
PT New polypeptide, useful for identification of Physcomitrella
PT for modulating stress resistance of a plant, comprises an
PT transcription factor stress-related protein derived from P
PT patens.
XX
PS Claim 13, Fig 3D, 129pp; English.
XX
XX
CC The invention relates to an isolated transcription factor;
CC protein (TFSRP) (I) selected from an AP2 similar-2 protein
CC zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 pr
CC protein, a MYB-1 protein (MYB-1), a CAAAT-box binding facto
CC (CABF-3), a sigma factor like protein (SFL-1) and its orth
CC described is an isolated TFSRP coding nucleic acid (II) wh
CC (I); and an isolated recombinant expression vector (III) c
CC where expression of (III) in a host cell results in increa
CC to environmental stress as compared to a wild type variety
CC cell. (I) or (II) is useful as markers for specific region
CC genome. (I) or (II) is useful for identification of Physco
CC and related organisms, mapping of genomes of P. patens
CC patens, identification and localisation of P. patens sequ
CC interest, evolutionary studies, determination of TFSRP reg
CC for function, modulation of a TFSRP activity, modulation o
CC metabolism of one or more cell functions, modulation of th
CC transport of one or more compounds and modulation of stres
CC (II) is useful for generating probes and primers for ident
CC cloning TFSRP homologues in other cell types and other org
CC as TFSRP homologues from mosses and related species, and f
CC and protein structural studies. AAU71797-AAU71804 represen
CC TFSRP amino acid sequences of the invention
XX
XX
SQ Sequence 188 AA;

Query Match 100.0%; Score 984; DB 5; Length 18
Best Local Similarity 100.0%; Pred. No. 6.9e-92; Matches 188; Conservative 0; Mismatches 0; Indels 0;

QY 1 MATERVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLGSKCYRETYWQAK 60
DB 1 MATERVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLGSKCYRETYWQAK 60

Best Local Similarity 42.7%; Pred. No. 4,8e-34;
Matches 82; Conservative 28; Mismatches 47; Indels 35; Gaps 3;
QY 8 QETTSQABEGPVMCKNLGFGSGQATMGLCSKCYRETYMOKMTALAQTA----- 60
DB 4 KEACGQOEGPILCINNCGFPGSAATMNMCSKCHKEMIKERQAKLAASSIYNVCGDG 63
QY 61 -----AQTSAATAAVQPPAPVHETKLTCEVERTMIVPHQSSSYQODLVTPAAAPQAVK 115
DB 64 GKEHTVAASGSTAVAN-----AQVEAKTLV-----VQPTDVAGTSEE 100
QY 116 SSIAPSRPEPNRCGSCRKRVGLTGFKRCGNLYCALHRSYDKHTCTYDYKAAQGEALAK 175
DB 101 VAVVKKVKEGPRKCATCRKRVGLTGFCNCGNMVYCALHRSYDKHTCQPDYRTAAADALAK 160
QY 176 ANPLVAEKVVK 187
DB 161 ANPVVKAEXLDDK 172
RESULT 5
ADY97065
ID ADX97065 standard; protein; 179 AA.
AC ADX97065;
XX
DT 21-APR-2005 (first entry)
DB plant full length insert polypeptide seqid 59729.
QY plant protectant; plant growth regulant; gene therapy; plant;
DB recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX Unidentified.
OS
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
XX
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LITUJ/) LITU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX WPI, 2004-180133/17.
XX
DR
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 59729; 15bp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.npsco.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as

CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombinant in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yield
CC content. This is the amino acid sequence of a plant full l
CC polypeptide that can be used in the recombinant DNA constr
CC invention.
XX
SQ Sequence 179 AA;
QY
Query Match 41.9%; Score 412.5; DB 8; Length 1
Best Local Similarity 42.3%; Pred. No. 1.3e-33;
Matches 80; Conservative 28; Mismatches 50; Indels
QY 8 QETTSQABEGPVMCKNLGFGSGQATMGLCSKCYRETYMOKMTALAAC 60
DB 12 KEACGQOEGPILCINNCGFPGSAATMNMCSKCHKEMIKERQAKLAAS 71
QY 61 --AQTSAATAAVQPPAPVHETKLTCEVERTMIVPHQSSSYQODLVTPA 118
DB 72 GKGPALIAITGVAVP-----QVEEKTIA-----VQPK 109
QY 119 AAPSRPEPNRCGSCRKRVGLTGFKRCGNLYCALHRSYDKHTCTYDYKA 178
DB 110 IAKKEGPNRCATCRKRVGLTGFCNCGNTYCSMHRYSDKHDCQFDYRT 169
QY 179 LVVAEKVVK 187
DB 170 VVKAEXLDDK 178
RESULT 6
ADY09608
ID ADY09608 standard; protein; 179 AA.
XX
AC ADY09608;
XX
DT 21-APR-2005 (first entry)
XX
DB plant full length insert polypeptide seqid 65423.
XX
QY plant protectant; plant growth regulant; gene therapy; pla
DB recombinant DNA construct; physical array; plant breeding
XX cold tolerance; heat tolerance; drought tolerance; herbici
KW extreme osmotic condition; pathogen tolerance; pest tolera
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth
KW yield; plant growth; plant development; seed oil; protein
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
XX
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LITUJ/) LITU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAA/) TABASKA J E.
PA (CAOY/) CAO Y.

XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI, 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1; SEQ ID NO 65423; 15bp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034889. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomanan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 179 AA;
SQ
Query Match 41.9%; Score 412.5; DB 8; Length 179;
Best Local Similarity 42.3%; Pred. No. 1.3e-33;
Matches 80; Conservative 28; Mismatches 50; Indels 31; Gaps 3;
QY 8 QETTGAPEGPVWCKNLGFFGSOATMGLCSKYRETVWQAKKTLALAEATQA----- 60
DB 12 KEAGCGQRPGLCTINCGFSGAATMNNCSKCHKEMIKQBOAOLAAISIDSVNGDN 71
QY 61 --AQTATAPAAVOPAPVPHETKLTCEVERTMI VPHQSSSYQODLVTPAAAAPQAVKSI 118
DB 72 GNGPAAATVGVAVP-----QVEKTI-----VQPMVAETSAALV 109
QY 119 AAPSPREPNCSSCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDKAAGGALAKNP 178
DB 110 IAKAKGFPKRCATCRKRVGLTGFNCGCGNTYCSMRHYSKDKHCQGFYRRAADALAKNP 169
QY 179 LVVAEKVVK 187
DB 170 VVKAERLKD 178

RESULT 7
ABM73978
ID ABM73978 standard; protein; 169 AA.
XX
XX ABM73978;
XX
XX 17-OCT-2003 (first entry)
XX
XX DNA clone ordinating in barley containing SNP sequence #388.
XX
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
XX
XX Hordeum vulgare.
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
PF

XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
XX (UTNI-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties;
PT sequences containing them for analysis and identification
PT varieties and production of barley transformants with des-
XX characteristics.
XX
XX Disclosure; SEQ ID XX; 284bp; Japanese.
XX
XX The present invention relates to oligonucleotide clones of
CC barley (Hordeum vulgare) which contain single nucleotide
CC (SNP). The oligonucleotides may be used for analysis of SN
CC varieties, identification of particular varieties and gen
CC analysis, isolation of specific genes and creation of new
CC transformation of barley varieties with them and producti
CC varieties with desired properties. The present sequence is
CC oligonucleotide clone sequence featured in the specificat
CC sequence data for this patent did not form part of the pri
CC specification, but was obtained in electronic format direc
CC at ftp.wipo.int/pub/published-pct-sequences
XX
XX Sequence 169 AA;
SQ
Query Match 41.5%; Score 408.5; DB 7; Length 1
Best Local Similarity 45.3%; Pred. No. 3e-33;
Matches 86; Conservative 25; Mismatches 44; Indels 5;
QY 8 QETTGAPEGPVWCKNLGFFGSOATMGLCSKYRETVWQAKKTLALAEATQA----- 60
DB 4 KETGCGSREGPILCTVNNCGFFGSAATMNNCSKCHKEMAMKQOALAA 63
QY 58 TQAAATATAAVOPAPVPHETKLTCEVERTMI VPHQSSSYQODLVTP 117
DB 64 VKEHLAAGSTAVAV-----AHVQAKTLITPA-----DIAGI 106
QY 118 IAAAPSPREPNCSSCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDY 177
DB 107 -----PSRCSTKRVGLTGFNCGCGNTYCATHRYSKDKCKFDY 158
QY 178 LVVAEKVVK 187
DB 159 PVVKAERLKD 168

RESULT 8
ADX75512
ID ADX75512 standard; protein; 173 AA.
XX
XX ADX75512;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polypeptide seqid 44878.
XX
XX plant proectant; plant growth regulant; gene therapy; pla
XX recombinant DNA construct; physical array; plant breeding
XX cold tolerance; heat tolerance; drought tolerance; herbic
XX extreme osmotic condition; pathogen tolerance; pest tolera
XX growth rate; cell cycle pathway; disease resistance;
XX galactomanan production; lignin production; plant growth
XX yield; plant growth; plant development; seed oil; protein
XX protein content.

Db 6 KEAGCQOPEGPILCINNCGFSGSAATMNMCSKCHKE-----MITKQEOAQLAASSISDI 59
Qy 67 -----TAAAVOPAPVHETKLTCEVERTWIVPHQSSSYQODLVTAAAAPOAVK 115
Db 60 VNGGDGKGKPVLAASVNAVP-----QVEQKTIIV-----YQPLVAVETSBA 100
Qy 116 SSIAAPSRPEPRRCGSKRKGVLGTGPKRCGNLYCALHRYSDKHTCTYTKAAGGELATK 175
Db 101 AAVIPKAKEGPDRCACRKRKVLGTGFSRCGNMYSVHRYSDBKHQCFDRTAADALAK 160
Qy 176 ANPLVVAEKVK 187
Db 161 ANPVVRAEKLDK 172
RESULT 10
ADX95012
ID ADX95012 standard; protein; 173 AA.
AC ADX95012;
DT 21-APR-2005 (first entry)
DE Plant full length insert polypeptide seqid 57676.
KW plant proectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX Unidentified.
XX OS
XX PN US2004034888-A1.
XX PD 19-FEB-2004.
XX PF 28-APR-2003; 2003US-00425114.
XX PR 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX PA (LIUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVAV/) KOVALIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABAKA/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WP; 2004-180133/17.
XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX PT pests, for conferring increased resistance to plant disease, or for
XX PT improving yield.
XX PS Claim 1; SEQ ID NO 57676; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX CC available in electronic form from the US patent office at
XX CC ftp.secddata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
XX CC of the invention are also useful in physical arrays of molecules and as
XX CC plant breeding markers. The recombinant DNA construct is useful for
XX CC improving plant tolerance to cold, heat, drought, herbicides, extreme
XX CC osmotic conditions, pathogens or pests, for manipulating growth rate in
XX CC plant cells by modification of the cell cycle pathway, for conferring

CC increased resistance to plant disease, for producing galat
CC lignin or plant growth regulators, for increasing the rat
CC recombination in plants, for improving yield by modificat
CC photosynthesis or carbohydrate, nitrogen or phosphorus us
CC or by providing improved plant growth and development und
CC stress condition or for modifying seed oil or protein yie
CC content. This is the amino acid sequence of a plant full
CC polypeptide that can be used in the recombinant DNA const
CC invention.
CC XX
SQ Sequence 173 AA;
Query Match 40.8%; Score 401.5; DB 8; Length 1
Best Local Similarity 42.7%; Pred. No. 1.6e-32;
Matches 82; Conservative 28; Mismatches 45; Indels 4;
Qy 8 OETTSQAPGEPVMCKNLCGFFGSOATMGLGSKCYRETWQAKMTALAK
Db 6 KEAGCQOPEGPILCINNCGFSGSAATMNMCSKCHKE-----MITKQE
Qy 67 -----TAAAVOPAPVHETKLTCEVERTWIVPHQSSSYQODL
Db 60 VNGGDGKGKPVLAASVNAVP-----QVEQKTIIV-----
Qy 116 SSIAAPSRPEPRRCGSKRKGVLGTGPKRCGNLYCALHRYSDKHTCTY
Db 101 AAVIPKAKEGPDRCACRKRKVLGTGFSRCGNMYSVHRYSDBKHQCF
Qy 176 ANPLVVAEKVK 187
Db 161 ANPVVRAEKLDK 172
RESULT 11
ABM90052
ID ABM90052 standard; protein; 160 AA.
XX AC ABM90052;
XX DT 02-JUN-2005 (first entry)
XX DE Rice abiotic stress responsive polypeptide SEQ ID NO:8389
XX KW abiotic stress tolerance; transgenic plant; plant; cereal
XX OS Oryza sativa.
XX OS WO2003008540-A2.
XX PN 30-JAN-2003.
XX PD 21-JUN-2002; 2002WO-US019668.
XX PF 22-JUN-2001; 2001US-0300112P.
XX PR 24-AUG-2001; 2001US-0314662P.
XX PR 26-SEP-2001; 2001US-0325277P.
XX PR 21-NOV-2001; 2001US-0332132P.
XX XX
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Krepis J, Briggs SP, Cooper B, Glazebrook J, Goff SA,
XX PI Moughamer T, Provart N, Rieke D, Zhu T;
XX DR WP; 2003-248011/24.
XX XX
XX PT New stress-responsive nucleic acid, useful for altering t
XX PT responsiveness of a plant, e.g. cereal, to an abiotic stre
XX PT stress, salt stress or osmotic stress.
XX XX
XX PS Claim 1; SEQ ID NO 8389; 89pp; English.
XX CC The invention relates to novel abiotic stress responsive
XX CC and polypeptides. Also disclosed are vectors, expression

PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-013149P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134421P.
PR 14-MAY-1999; 99US-013470P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135529P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145213P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148117P.
PR 11-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153707P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158233P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P-
PR 14-OCT-1999; 99US-0159638P-
PR 18-OCT-1999; 99US-0159584P-
PR 21-OCT-1999; 99US-0160741P-
PR 21-OCT-1999; 99US-0160767P-
PR 21-OCT-1999; 99US-0160768P-
PR 21-OCT-1999; 99US-0160770P-
PR 21-OCT-1999; 99US-0160814P-
PR 21-OCT-1999; 99US-0160815P-
PR 22-OCT-1999; 99US-0160980P-
PR 22-OCT-1999; 99US-0160981P-
PR 22-OCT-1999; 99US-0160989P-
PR 25-OCT-1999; 99US-0161404P-
PR 25-OCT-1999; 99US-0161405P-
PR 25-OCT-1999; 99US-0161359P-
PR 26-OCT-1999; 99US-0161359P-
PR 26-OCT-1999; 99US-0161360P-
PR 26-OCT-1999; 99US-0161361P-
PR 28-OCT-1999; 99US-0161920P-
PR 28-OCT-1999; 99US-0161992P-
PR 28-OCT-1999; 99US-0161993P-
PR 29-OCT-1999; 99US-0162142P-
PR 29-OCT-1999; 99US-0162142P-

Query Match 38.5%; Score 379; DB 3; Length 173;
Best Local Similarity 43.5%; Pred. No. 3, 2e-30;
Matches 77; Conservative 34; Mismatches 46; Indels 20; Gaps 5;

QY 15 PEGPWCKKLCGFFSGQATMGCSCKYRETWQAKMTALAEQTAAQNTSTAALVQPP 74
DB 12 PEGPKLTCTNCCGFFGSAATMNNCKCHKDMLPQ-----QEGAKFASVSGTSSS----- 61
QY 75 AAVHETKLTCEVERMTIVPHOSSYQODLVTPAALAPQAVKSIAP---SRPE-PNRCG 130
DB 62 -----SNIKETFTALVDIETKSVPMYVSQPSVQV- AEVVAPEBAAPKPKPSCT 115
QY 131 SCRKRVGLTGFCRCGNTLYCALHRYSDKATCTYDYKAAGQEAIAKANPLVAEKVVK 187
DB 116 TENKRVGLTGFCRCGSLFCGTHRYADVHDCSFNHAAGEAIAKANPLVNAEKLDK 172

RESULT 14

AAG49664
AAG49664 standard; protein; 173 AA.

AC AAG49664;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 62852.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P-
PR 05-MAR-1999; 99US-0122180P-
PR 09-MAR-1999; 99US-0123548P-
PR 23-MAR-1999; 99US-0125788P-
PR 25-MAR-1999; 99US-0126264P-
PR 29-MAR-1999; 99US-0126785P-
PR 01-APR-1999; 99US-0127462P-
PR 06-APR-1999; 99US-0128234P-
PR 08-APR-1999; 99US-0128714P-
PR 16-APR-1999; 99US-0129845P-
PR 19-APR-1999; 99US-0130077P-

[illegible]

PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	22-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	07-MAY-1999;	99US-0132487P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137528P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140639P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144532P.
PR	20-JUL-1999;	99US-0144584P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145154P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	03-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	03-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	06-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 38.5%; Score 379; DB 3; Length 173;
Best Local Similarity 43.5%; Pred. No. 3.2e-30;
Matches 77; Conservative 34; Mismatches 46; Indels 20; Gaps 5;

Qy 15 PEGPVWCNULCGPFGSOATWGLGSKCYRETVNQAKMTALAEQATOAQATSAATAVQPP 74
Db 12 PEGPKLCTNNGCFGSAATNMCSKCHKDWLFO-----OEGAKFASAVSGTSS---- 61
Qy 75 APVHETKLTCEVERTMIVPHQSSSYQODLVTPAAAPOAVKSSIAAP---SRPE-PNRCG 130
Db 62 ----SNIKETFTALVDIETKSEVPMTVSQPSVQVV-AEVVAPEEAKPKGSRCT 115
Qy 131 SCRKRVLGTGPKRCGNLYCALHRYSDKTCYDYKAGOEALAKANPLVNAEKVYK 187
Db 116 TCNKRVLGTGPKRCGSLFCGTHRYADVHDCSFNYHAAQEAALAKANPVVKAEXLDK 172

Search completed: December 8, 2005, 16:29:10
Job time : 137 secs

This Page Blank (uspto)

A/Reference number: 225013
A/Accession: T49033
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-170 <PUN>
A/Cross-references: UNIPROT:O9LX15; UNIPARC:UPI000009DQFD; EMBL:AL353912; GSPDB:GN00061
A/Experimental source: cultivar Columbia; BAC clone F3C22
C/Genetics:
A/Gene: ATSP:F3C22.200
A/Map position: 3

Query Match 36.8%; Score 362.5; DB 2; Length 170;
Best Local Similarity 37.8%; Pred. No. 2.9e-23;
Matches 73; Conservative 32; Mismatches 53; Indels 35; Gaps 4;

6 VSOETTSQAPRGPWCKNLGFFGSOATMGLCSKCYRETVMQAKMTALAEQATOAQATS 65
1 MAEHRCOTPEPSNRLCVNNCGFLSSATMGLCSNCTGDLCLK-----QQQSSISSTV 54

66 ATAAAVOPRAPHETKLTCEVERTMIVPHQSSYQODLVTPAAAAPQ-----VKSIIA 119
55 ESSLSVSPPS-----SSSSSISSPIPLLNKPSVKLEVEPEKKAIVS 96

120 APS-----RPPNRCGSRKRVGLTGFKRCGNLYCALHRSDKGTCTYDPAKAGQELA 174
97 LPTTEONQOQPNRCCTCRKRVGLTGFKRCGTCVHRYPEIHGCSYDPKSAGREBIA 156

175 KANPLVVAEKVVK 187
157 KANPLVVAKAKLQK 169

RESULT 3
TS1098
hypochemical protein p85RF [imported] - Prunus armeniaca
C/Species: Prunus armeniaca
C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C/Accession: TS1098
R/Mbegule-A-Mbegule, D.; File-Lycaon, B.R.
submitted to the EMBL data library, March 1999
A/Description: Molecular cloning and nucleotide sequence of a putative protein from apricot
A/Reference number: Z55296
A/Accession: TS1098
A/Status: Preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-173 <MBE>
A/Cross-references: UNIPROT:O9XH71; UNIPARC:UPI00000AC882; EMBL:AF139499; PTDN:AAD38146
C/Genetics:
A/Note: p85RF

Query Match 36.7%; Score 361.5; DB 2; Length 173;
Best Local Similarity 39.2%; Pred. No. 3.6e-23;
Matches 76; Conservative 26; Mismatches 53; Indels 39; Gaps 5;

8 QETTSQA-PEGPVMCKNLGFFGSOATMGLCSKCYRETVMQAKMTALAEQ-----TQ 59
4 EETGQGPHEGPILCVNNCGFFGSVATRMGSKCHKMDLKEBQAKLAASFGNIVNGTS 63

60 AAQATSATAAVOPRAPHETKLTCEVERTMIVPHQSSYQODLVTPAAAAPQ-----AV 114
64 NSNGNEPVAA-----GVDVDAHLVEPPTISLQPSFSGS 98

115 KSTIAAPRPE-PNRCGSRKRVGLTGFKRCGNLYCALHRSDKGTCTYDPAKAGQRAI 173
99 GSGSGEAKPEGPKRKCGTCNKRVLGTGFCNCGHLFCVHRYSDKHDCPYDHTAARDVI 158

174 KANPLVVAEKVVK 187
159 KANPLVVAKADLQK 172

RESULT 4
T006611

```

hypothetical protein F16J13.110 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
C:Accession: T06611
R:Bayan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Du
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15789
A:Accession: T06611
A:Molecule type: DNA
A:Residues: 1-175 <BEV>
A:Cross-references: UNIPROT:Q9S269; UNIPARC:UPI00000A7C92; EMBL
A:Experimental source: cultivar Columbia; BAC clone F16J13
C:Genetics:
A:Gene: ATSP:F16J13.110
A:Map position: 4

Query Match          36.7%  Score 361.5; DB 2; Length 1
Best Local Similarity 41.0%  Pred. No. 3,66-23;
Matches 73; Conservative 29; Mismatches 61; Indels

QY      10  TTGAAPGAPVWCKNLGSPFGSOATMGLCSKCYRETWQAQMTALAEOAT
      69  :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      8  STSPPTPEPKLCDNGCGPFGSPSMNLCSKCYRSLRAEBDQTAAKAAV
      65  :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      70  AVQPPAPVHETKLTCEVERMTVPHQSSSYQODLVTPAALPAQVKS:
      127  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      66  SIAPGQKHPIEKPAHLETVTVTAEPSS-----VPVAAEODE
      114  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      128  RCGSCRKRVGLTGFKCGCNLYCALHRYSDKHCTVYKKAAGCAIYKA
      5   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      115  RCFCKNKKVGVWGFCKCGSTFCGSHYPEKHBCSFPDKVEDAIYKA
      2   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 5
CG4779
hypothetical protein At2g36320 [imported] - Arabidopsis thalian
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change
C:Accession: CG4779
R:Li, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.;
M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.F
euss, D.; Nierman, W.C.; White, O.; Eiden, J.A.; Salzberg, S.L.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Ara
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: CG4779
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <STO>
A:Cross-references: UNIPROT:Q9SJM6; UNIPARC:UPI00000A045C; GB:A
C:Genetics:
A:Gene: At2g36320
A:Map position: 2

Query Match          36.2%  Score 356; DB 2; Length 161
Best Local Similarity 39.8%  Pred. No. 9,66-23;
Matches 74; Conservative 29; Mismatches 53; Indels

QY      6  VSQRTTGAAPGAPVWCKNLGSPFGSOATMGLCSKCYRETWQAQMTALA
      65  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1  MAEHRRCETPEGHRLCVNNCGPFGSSATMNLCSNYCDLCLKQO-----
      55  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      66  ATAAAVOPAPVHETKLTCEVERMTVPHQSSSYQODLVTPAALPAQ--
      122  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      56  SLSPVT--APVLE-----NTAAELBIPPTYKTEBK
      94  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      123  RPE-PNRGSCRRKRVGLTGFKRCGNLYCALHRYSDKHCTTYDYKAAQ
      181  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      95  PPGAPNRCTVGRKRVGLTGFKRCGTTFCGSHRIPVYHGCITDFPKSAGR
      154  ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      182  AEKVVK 187
      181  |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      155  AAKLQK 160
      154  |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```


RESULT 6
D84674
Hypothetical protein At2g27580 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84674
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bent, M.I.; Tom, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Accession: D84674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-163 <STO>
A:Cross-references: UNIPROT:Q9ZNU9; UNIPARC:UPI00000AC021; GB:AE002093; NID:G3860274; PI
C:Genetics:
A:Gene: At2g27580
A:Map position: 2

Query Match 35.2%; Score 346.5; DB 2; Length 163;
Best Local Similarity 38.2%; Pred. No. 6e-22;
Matches 76; Conservative 30; Mismatches 44; Indels 49; Gaps 6;

QY 1 MATERVSETTQAPBGPVMCKNLGFGSSQATMGLCSKCYRETYMQAKMTALAQA-TQ 59
DB 1 MAEEHRLQ-----PRLCANNGCFPGSTATQNLCSKCPDLQHQNSSSTAKHALTQ 52
QY 60 AAQATSAT-AAVOPAPVHETKLTCEVERMTIVPHOSSYQODLVTPAAAPQAV--KS 116
DB 53 SLAAAGAAASSSVSPPP-----PPADKEIVEAKS 83
QY 117 STAFSRPE-----PARCGSKRVGLTGFKRCGNLYCALHRYSDKHTCTDYKAA 168
DB 84 EGRBAAEPEAGPPDPDRCLTCRRRVGITGFRCKGVPFGCTRHYAQCHECSDFPKM 143
QY 169 GQBAIAKANPLVAAEKVK 187
DB 144 GKDIAKANPIVADLEK 162

RESULT 7
T04577
Hypothetical protein T12H17.210 - Arabidopsis thaliana
N:Alternate names: hypothetical protein F7H19.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
R:Bevan, M.; Hilbert, H.; Brun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, A.; Reference number: 215377
A:Accession: T04577
A:Molecule type: DNA
A:Residues: 1-176 <BEV>
A:Cross-references: UNIPROT:Q49663; UNIPARC:UPI00000ABF74; EMBL:AL021635
A:Experimental source: cultivar Columbia; BAC clone T12H17
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirks, W.; Stiekema, W.; Bancroft, I.; Mew submitted to the Protein Sequence Database, February 1998
A:Reference number: 215377
A:Accession: T04577
A:Molecule type: DNA
A:Residues: 1-176 <BEV>
A:Cross-references: UNIPROT:Q49663; UNIPARC:UPI00000ABF74; EMBL:AL021635
A:Experimental source: cultivar Columbia; BAC clone T12H17
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirks, W.; Stiekema, W.; Bancroft, I.; Mew submitted to the Protein Sequence Database, July 1998
A:Reference number: 215399
A:Accession: T05114
A:Molecule type: DNA
A:Residues: 1-176 <BEW>
A:Cross-references: UNIPARC:UPI00000ABF74; EMBL:AL031018
C:Genetics:
A:Map position: 4
A:Note: T12H17.210; F7H19.10

Query Match 34.3%; Score 338; DB 2; Length 176;
Best Local Similarity 41.2%; Pred. No. 3.3e-21;

Matches 75; Conservative 29; Mismatches 58; Indels 6;

QY 10 TTQAPBGPVMCKNLGFGSSQATMGLCSKCYRETYMQAKMTALAEQ 69
DB 12 TQSQASE--PKLCVKGCGFGSPSNDLCSKCYRGI CAEBAOTAVAKAA 70
QY 70 --AVOPAPVHETKLTCEVERMTIVPHOSSYQODLVTPAAAPQAVK 126
DB 71 LFIAPPAVVEPK--PEAAAVVVAEPPS-----SAVPEANE 114
QY 127 NRCGSKRVGLTGFKRCGNLYCALHRYSDKHTCTDYKAGQBAIA 186
DB 115 NRCLCNKKVGMGRKCKGSPFCGEHRYPERHDCSPDKVGRGELAH 174
QY 187 KF 188
DB 175 RF 176

RESULT 8
T21254
Hypothetical protein F22D6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change
C:Accession: T21254
R:Wilkinson, J. submitted to the EMBL Data Library, April 1996
A:Reference number: 219397
A:Accession: T21254
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-189 <WIL>
A:Cross-references: UNIPROT:Q19723; UNIPARC:UPI000007F768; EMBL
A:Experimental source: clone F22D6
C:Genetics:
A:Gene: CESP:F22D6.2
A:Map position: 1
A:Introns: 51/1, 106/2, 141/1

Query Match 31.8%; Score 313; DB 2; Length 18;
Best Local Similarity 35.4%; Pred. No. 4.3e-19;
Matches 68; Conservative 31; Mismatches 69; Indels 6;

QY 8 QETTSQAPBGPVMCKNLGFGSSQATMGLCSKCYRETYMQAKMTALAEK 67
DB 5 QOQAOCTAPS---CRAGCGFGASATBEGYCSQCFKNTILKROODTVRLT 60
QY 68 AAAYG-PPAPV-----HETKLTCEVERMTIVPHOSSYQOD 113
DB 61 SSALSKSPSSVDMCKAAVSVSDETAKMDCR---DIINVCDDQINDR 115
QY 114 VKSSIAAPSRPEPNRCGSKRVGLTGFKRCGNLYCALHRYSDKHTCT 173
DB 116 ITVDVPPVVK--KANRCHMKRVRGLTGRSCRCGGLYCGDHRHYDQAHNC 174
QY 174 AKANPLVAAEKV 185
DB 175 KKNPVPVVDK 186

RESULT 9
T11846
pathogenesis-related protein 3 - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change
C:Accession: T11846
R:Sharma, Y.K.; Hinojos, C.M.; Mehdy, M. Mol. Plant Microbe Interact. 5, 89-95, 1992
A:Title: CDNA cloning, structure and expression of a novel pat
A:Reference number: 217354; MUID:92288407; PMID:1600239
A:Accession: T11846
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

lated

```

A:Residues: 1-117 <SHA>
A:Cross-references: UNIPROT:Q41123, UNIPARC:UPI00000A4E14, EMBL:M75856, NID:G165362, PDB:
C:Genetics:
A:Gene: PR3

Query Match      28.5%; Score 280; DB 2; Length 137;
Best Local Similarity 42.9%; Pred. No. 1.8e-16;
Matches 57; Conservative 21; Mismatches 43; Indels 12; Gaps 3;

Qy 55 EQATQAQATSATMAAVOPPAVHEHTKLTCEVERTMIVPHQSSYQQDLVTPAALPAQAV 114
Db 16 EPTPCATATAAATSISEPSRFPDAPATTSRS---PKRS-----LPLEDAAND 64

Qy 115 KSIAPSPREPPRCGSCRKVGLTGFCRCGNLYCALHRSYDKTCTDYAAAGEALA 174
Db 65 RTVAASEPKR-AVNRCSGCRRRVGLTGFRRCGDLCAEHRVYDRHDCSYKTVGREALA 123

Qy 175 KANPLVAEKVKV 187
Db 124 RENPVAKAKIVK 136

RESULT 10
T10218
[hypothetical protein T30C3.50 - Arabidopsis thaliana]
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16897
A:Accession: T10218
A:Cross-references: UNIPROT:Q9STU9, UNIPARC:UPI0000A0710, EMBL:AL079350, GSPDB:GN00062,
A:Experimental source: cultivar Columbia; BAC clone T30C3
C:Genetics:
A:Gene: ATSP:T30C3.50
A:Map position: 4

Query Match      26.0%; Score 255.5; DB 2; Length 130;
Best Local Similarity 28.0%; Pred. No. 1.9e-14;
Matches 47; Conservative 30; Mismatches 44; Indels 47; Gaps 1;

Qy 21 CKNLGCFFGQATWGLCSKCYRETVMQAKMTALAQATQAQATSAVOPPAVHEHT 80
Db 10 CEGGCGLGYGRVNNNLSLCYKKSIVLQHSPLRFPETEGSQCCPPTNSPAVEEPEV--- 66

Qy 81 KLTCEVERTMIVPHQSSYQQDLVTPAALPAQAVKSSIAFSPREPPRCGSCRKVGLTG 140
Db 67 -----KRRRCGICRKVKMG 82

Qy 141 FKRCGNLYCALHRSYDKTCTDYKAAGEALAKANPLVAEKVKV 188
Db 83 FKRCGMFCGSHRYPEBHSCEPDYKSGRLALATQLPIRADKLORF 130

RESULT 11
JN0673
ubiquitin-like fusion protein Anla - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 05-Oct-2004
C:Accession: JN0673
R:Linmen, J.M.; Bailey, C.P.; Weeks, D.L.
Gene 128, 181-188, 1993
A:title: Two related localized mRNAs from Xenopus laevis encode ubiquitin-like fusion pr
A:Reference number: JN0673; MUID:93292985; PMID:8390387
A:Accession: JN0673
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-693 <LIN>
A:Cross-references: UNIPROT:Q91889, UNIPARC:UPI000017A486
C:Genetics:

```

[illegible]

C/Species: Bos primigenius taurus (cattle)
C/Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C/Accession: I47035
R/Shima, D.T.; Saunders, K.B.; Gougos, A.; D'Amore, P.A.
Differentiation 58, 217-226, 1995
A/Title: Alterations in gene expression associated with changes in the state of endothel
A/Reference number: I47034; MID:95228954; PMID:7713329
A/Accession: I47035
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-46 <SH1>
A/Cross-references: UNIPROT:Q28900; UNIPARC:UPI000006679C; GB:S77733; NID:g998679; PIDN:

Query Match 18.0%; Score 177; DB 2; Length 46;
Best Local Similarity 69.8%; Pred. No. 2.4e-08;
Matches 30; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 143 CRCGNLYCALHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKV 185
Db 1 CRCGNLYCALHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKI 43

RESULT 14

T18125
hypothetical protein A623L - Chlorella virus PBCV-1
C/Species: Chlorella virus PBCV-1
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18125
R/Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A/Reference number: Z18806
A/Accession: T18125
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-67 <GRA>
A/Cross-references: UNIPROT:Q41105; UNIPARC:UPI0000066F20; EMBL:U42580; NID:g4028896; PI
A/Experimental source: specific host Chlorella strain NC64A
C/Genetics:
A/Note: A623L

Query Match 12.6%; Score 124; DB 2; Length 67;
Best Local Similarity 36.7%; Pred. No. 0.00092;
Matches 22; Conservative 10; Mismatches 18; Indels 10; Gaps 1;

Qy 127 NRCGCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKV 186
Db 10 SRCEICRKKTKGLGFVCKGCHTFCEKGRIMESHSC-----PTLQAKERIILEKV 59

RESULT 15

S43586
F26F3.4 protein (clone F26F3) - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S43586
R/Suleton, J.
submitted to the EMBL Data Library, April 1994
A/Reference number: S43582
A/Accession: S43586
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-341 <SUL>
A/Cross-references: UNIPROT:Q20899; UNIPARC:UPI00000829EB; EMBL:Z32681; NID:g474790; PID
C/Genetics:
A/Introns: 33/3; 101/2; 192/2; 232/3; 277/3

Query Match 11.4%; Score 112.5; DB 2; Length 341;
Best Local Similarity 28.6%; Pred. No. 0.041;
Matches 24; Conservative 12; Mismatches 33; Indels 15; Gaps 2;

Qy 112 QAVKSSIAAPSRPE-----PNRCGCRKRVGLT--GFKRCGNLYCALHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKV 186
Db 246 QCRKSNVLAIPSSBEELLENLKKFEARKTKCNTCFKKLSAAGQTMKCKLRIFCDRHRHP 305

Qy 157 DKHTCTYDYKAGQEAIAKANPLV 180
Db 306 KNHTCTYDYKQDGRNKKLKKGNNSKV 329

Search completed: December 8, 2005, 16:32:58
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 12:58:00 ; Search time 159 Seconds

(without alignments)
834.210 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSQETTSQARBPV.....GGEAIKAKNPLVAKKVKF 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705526306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	419.5	42.6	171	2	Q69WY4_ORYSA
2	416.5	42.3	171	2	Q6H7P8_ORYSA
3	415.5	42.2	171	2	Q6VAG2_ORYSA
4	413	42.0	233	2	Q6R318_MALZE
5	386	39.2	160	2	Q852K5_ORYSA
6	384.5	39.1	161	2	Q7Y1W9_ORYSA
7	384	39.0	173	2	Q8H0X0_ARATH
8	379	38.5	173	2	Q9SYC3_ARATH
9	369	37.5	211	2	Q6E1P3_XENLA
10	365	37.1	211	2	Q6DF90_XENLA
11	365	37.1	213	1	Z20D2_MOUSE
12	364	37.0	213	2	Q6PBR4_BRARE
13	364	37.0	213	2	Q6PBR4_BRARE
14	363.5	36.9	212	2	Q5ZUR4_CHICK
15	362.5	36.8	170	2	Q9LX15_ARATH
16	361.5	36.7	173	2	Q9XK71_PPRAR
17	361.5	36.7	175	1	ZF2N3_ARATH
18	361.5	36.7	212	2	Q7ZU97_BRARE
19	361	36.7	224	2	Q6Z541_ORYSA
20	356	36.2	161	1	ZF2N2_ARATH
21	352	35.8	169	2	Q94B40_ARATH
22	352	35.8	223	2	Q6DGF4_RAT
23	352	35.8	223	2	Q9DCH6_MOUSE
24	350.5	35.6	210	2	Q4G2S7_PORCU
25	347	35.3	611	2	Q4SD48_TERNG
26	346.5	35.2	163	2	Q9ZNU9_ARATH
27	346	35.2	207	2	Q4T4U6_TERNG
28	344.5	35.0	224	2	Q4T017_TERNG
29	343.5	34.9	169	2	Q852K6_ORYSA
30	343.5	34.9	208	2	Q9GZY3_HUMAN
31	343	34.9	170	2	Q84PD6_ORYSA

32	343	34.9	223	2	Q9ER79_MOUSE	Q9E
33	341.5	34.7	208	2	Q5R7S6_PONPY	Q5R
34	340	34.6	199	2	Q9VHF4_DROME	Q9V
35	338.5	34.4	148	2	Q942F8_ORYSA	Q94
36	338	34.3	168	1	ZF2N1_ARATH	Q6n
37	338	34.3	176	2	Q49663_ARATH	Q49
38	338	34.3	206	2	Q8W062_DROME	Q8W
39	337.5	34.3	164	1	ISAP1_ORYSA	Q91
40	337.5	34.3	202	2	Q96038_CIOGA	Q96
41	335.5	34.1	198	2	Q7QAK6_XANGA	Q7Q
42	334	33.9	201	2	Q6P840_XENTR	Q6P
43	333.5	33.9	154	2	Q6H754_ORYSA	Q6H
44	332.5	33.8	208	2	Q9BQF7_HUMAN	Q9B
45	327.5	33.3	208	2	Q6T1P0_HUMAN	Q6T

ALIGNMENTS

RESULT 1	
Q69WY4_ORYSA	
ID Q69WY4_ORYSA PRELIMINARY; PRT; 171 AA.	
AC Q69WY4;	
DT 25-OCT-2004 (TREMBLrel. 28, Created)	
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)	
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)	
DE Putative multiple stress-responsive zinc-finger protein.	
GN Name=P0417G12.11; Synonyms=P0429G06.28;	
OS Oryza sativa (japonica cultivar-group).	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trac	
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae	
OC Ehrhartoideae; Oryzeae; Oryza.	
OX NCBI_TaxID=39947;	
RN (1)	
RP NUCLEOTIDE SEQUENCE.	
RA Sasaki T., Matsunoto T., Yamamoto K.;	
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, 1	
RT clone: P0417G12."	
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.	
RP NUCLEOTIDE SEQUENCE.	
RA Sasaki T., Matsunoto T., Yamamoto K.;	
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, 1	
RT clone: P0429G06."	
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AP003626; BAD35521.1; -; Genomic DNA.	
DR EMBL; AP003711; BAD3553.1; -; Genomic DNA.	
DR Gramene; Q69WY4; -;	
DR GO; GO:0003677; F:DNA binding; IEA.	
DR GO; GO:0008270; F:zinc ion binding; IEA.	
DR InterPro; IPR002653; Znf_A20.	
DR InterPro; IPR000058; Znf_AN1.	
DR Pfam; PF01754; zf-A20; 1.	
DR Pfam; PF01428; zf-AN1; 1.	
DR SMART; SM00259; Znf_A20; 1.	
DR SMART; SM00154; Znf_AN1; 1.	
KW Zinc-finger.	
SQ SEQUENCE 171 AA; 18402 MW; 5E425E95741816C4 CRC64;	
Query Match 42.6%; Score 419.5; DB 2; Length	
Best Local Similarity 45.6%; Pred. No. 7.4e-27;	
Matches 82; Conservative 26; Mismatches 59; Indels	
QY 8 QETTSQABEGPVMCKNLGFFGSGATMGLCSKCYRETYWQAKMTALAE	67
DB 4 KETGCGQEGEGILICINNCGFFGSAATMMCKSCHKEMTMKQEQAKLA-	61
QY 68 AAAYVPPAPVETKLTCEVERTMTIVPHQSSSYQDDLTTPAAAAPQAVK	127
DB 62 DSGKEPPIIAGAAVAAVAAVEKTVL-----AQPASIIAGSSEG	110
QY 128 RCGSCRKRVGLTGFCKRCGNLYCALHRSYDKHTCTYDYKAAQGEALIAV	187

```

DB      111 RCTCRKRVGLTGFNCRGNLYCAMHRYSDKHDCQFDYRTAARDAIAKANPVVAKKLDK 170

RESULT 2
O6H7P8_ORYSA PRELIMINARY; PRT; 173 AA.
AC      O6H7P8_
DR      05-JUL-2004 (TREMBLrel. 27, Created)
DT      05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE      Putative zinc-finger protein.
GN      Name=OJ1225_F07.15;
OS      Oryza sativa (japonica cultivar-group).
OC      Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzeae; Oryza.
OX      NCBI_TaxID=39947;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Sasaki T., Matsumoto T., Yamamoto K.;
RL      Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AP004184; BAD25251.1; -; Genomic_DNA.
DR      Gramene; O6H7P8; -.
DR      GO; GO:0003677; F:DNA binding; IEA.
DR      GO; GO:0008270; F:zinc ion binding; IEA.
DR      InterPro; IPR002653; Znf_A20.
DR      InterPro; IPR000058; Znf_AN1.
DR      Pfam; PF01754; zf-A20; 1.
DR      Pfam; PF01428; zf-AN1; 1.
DR      SMART; SM00259; Znf_A20; 1.
DR      SMART; SM00154; Znf_AN1; 1.
KW      Zinc-finger.
SQ      SEQUENCE 173 AA; 18540 MW; 5B3BC00752359ADE CRC64;

Query Match      42.3%; Score 416.5; DB 2; Length 173;
Best Local Similarity 42.7%; Pred. No. 1.3e-26;
Matches 82; Conservative 28; Mismatches 47; Indels 35; Gaps 3;

QY      8 QETTSQAPBEGPVMCKNLGCGFGSQATMGLCSKCYRETVMQAKMTALAQAQA----- 60
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      4 KEAGCQQPEGPILICINNGCFGSAATMNMCKCHKEMIMKEQAKLAASIDSYNGCDG 63
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      61 -----AQTATAAVOPPAVPHETKLTCEVERTMIVPHOSSYQODLVTPAAAPQAVK 115
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      64 GKEHTVAASGSTRNAV-----AQVEAKTLV-----VQPTDVAGTSEE 100
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      116 SSIAAPSRPEPNRCSCKRRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAAGQEAIAK 175
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      101 VAVVPEKVEGPRCATCRKRVGLTGFNCRGNLYCALHRYSDKHDCQFDYRTAARDAIAK 160
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      176 ANPLVVAEKVKV 187
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      161 ANPVVAKKLDK 172
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
O6VAG2_ORYSA PRELIMINARY; PRT; 171 AA.
AC      O6VAG2;
DR      05-JUL-2004 (TREMBLrel. 27, Created)
DT      05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE      Putative zinc-finger protein.
OS      Oryza sativa (indica cultivar-group).
OC      Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzeae; Oryza.
OX      NCBI_TaxID=39946;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Vydeshi K., Gupta A.K.;
RL      Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

```

```

DR      EMBL; AY345599; AAC084334.1; -; mRNA.
DR      Gramene; O6VAG2; -.
DR      GO; GO:0003677; F:DNA binding; IEA.
DR      GO; GO:0008270; F:zinc ion binding; IEA.
DR      InterPro; IPR002653; Znf_A20.
DR      InterPro; IPR000058; Znf_AN1.
DR      Pfam; PF01754; zf-A20; 1.
DR      Pfam; PF01428; zf-AN1; 1.
DR      SMART; SM00259; Znf_A20; 1.
DR      SMART; SM00154; Znf_AN1; 1.
KW      Zinc-finger.
SQ      SEQUENCE 171 AA; 18401 MW; 07760A1F5410B448 CRC64;

Query Match      42.2%; Score 415.5; DB 2; Length 171;
Best Local Similarity 45.0%; Pred. No. 1.6e-26;
Matches 81; Conservative 27; Mismatches 59; Indels 2;

QY      8 QETTSQAPBEGPVMCKNLGCGFGSQATMGLCSKCYRETVMQAKMTALAEAC 67
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      4 KETGCGQPEGPILICINNGCFGSAATMNMCKCHKEMIMKEQAKLA-- 61
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      68 AAAVQPPAPVPHETKLTCEVERTMIVPHOSSYQODLVTPAAAPQAVKS 127
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      62 DSGKEPPIAGHAENVAVQVEVKTLV-----AQPAEIAGPSSEGV 110
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      128 RCGSCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAAGQEAIAK 187
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      111 RCTCRKRVGLTGFNCRGNLYCAMHRYSDKHDCQFDYRTAARDAIAK 170
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
O6R318_MAIZE PRELIMINARY; PRT; 233 AA.
AC      O6R318;
DR      05-JUL-2004 (TREMBLrel. 27, Created)
DT      05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE      Putative zinc finger protein Zmzf.
OS      Zea mays (Maize).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tract
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae.
OC      PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX      NCBI_TaxID=4577;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Zhang Z.-X., Tang W.-H., Tao Y.-S., Zheng Y.-L.;
RL      Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY515607; AAS00453.1; -; mRNA.
DR      GO; GO:0003677; F:DNA binding; IEA.
DR      GO; GO:0008270; F:zinc ion binding; IEA.
DR      InterPro; IPR002653; Znf_A20.
DR      InterPro; IPR000058; Znf_AN1.
DR      Pfam; PF01754; zf-A20; 1.
DR      Pfam; PF01428; zf-AN1; 1.
DR      SMART; SM00259; Znf_A20; 1.
DR      SMART; SM00154; Znf_AN1; 1.
KW      Zinc-finger.
SQ      SEQUENCE 233 AA; 25019 MW; 5561DAC1C1478C6 CRC64;

Query Match      42.0%; Score 413; DB 2; Length 233;
Best Local Similarity 44.8%; Pred. No. 3.5e-26;
Matches 81; Conservative 30; Mismatches 58; Indels 3;

QY      8 QETTSQAPBEGPVMCKNLGCGFGSQATMGLCSKCYRETVMQAKMTALAC 67
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      63 KEAGCQPEGPILICINNGCFGSAATMNMCKCHKEMITTKDDQAKLA-- 120
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      68 AAAVQPPAPVPHETKLTCEVERTMIVPHOSSYQODLVTPA-AAAPQAVK 126
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      121 DAVMEPGS-----AGSTTVAVAVQVELQKKGVQPPAVALEPNEG 171
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      127 NRCSCKRRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAAGQEAIAK 186
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      172 NRCSCKRRVGLTGFNCRGNLYCALHRYSDKHDCQFDYRTAARDAIAK 231
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

OY 187 K 187
DB 232 K 232

RESULT 5

O852K5_ORYSA PRELIMINARY; PRT; 160 AA.
AC O852K5;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Putative zinc finger protein (putative zinc finger transcription factor ZFP38)
DE Name:OSUNB0060J21.18; Synonyms=ZFP38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaraloideae; Oryzaceae; Oryza.
NCBI_Taxid=39947;
RX
RN
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M., Overton I.L., Tselirin T., Kim M.M., Bera J.J., Jin S.S., Fadrosch D.W., Tallon L.J., Koo H., Ziemann V., Hsieh J., Blunt S., Vanaken S.S., Riedmiller S.B., Uterback T.T., Reichlyum T.V., Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RN Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
[3]
RN
RP NUCLEOTIDE SEQUENCE.
RA Huang J., Zhang H.-S.;
RN Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC090871; AA037974.1; -; genomic_DNA.
DR EMBL; AY377427; AA083587.1; -; mRNA.
DR Gramene; O852K5; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01428; Zf-A20; 1.
DR Pfam; PF01428; Zf-AN1; 1.
SQ SEQUENCE 160 AA; 17532 MW; 1A10D44CF546B232 CRC64;

Query Match 39.2%; Score 386; DB 2; Length 160;

Best Local Similarity 42.2%; Pred. No. 4e-24; Matches 81; Conservative 23; Mismatches 50; Indels 38; Gaps 5;

OY 1 MATE---RVSOETTSQAPGPGVMCKNLGCFGSGQATMGLGCKCYRETVMQAQMTALAQA 57
DB 1 MAQESWKKEAEETGHTPEAPILCVNCGFFSGSMTNNKSCYRDFV-KAK----- 51
OY 58 TQAQATSATAAAVQPPAPVHETKLTCEVERTMIIVPHOSSSYQODLVTPAAAPQAVKSS 117
DB 52 -----TVATVVEKKPLASLSTPLVTE-----VTDGSGSVADKQ 87
OY 118 IAAPEPRP--NRGSSCKRKGVLTFKRCRCGNLYCALHRSYDKTCTDYRAAGCEALAK 175
DB 88 VMEEDTPEPPSRKCLSCRRKQGLTFKRCRCGTFCSMRHYADSHKCTFDYKQVGREQIAK 147
OY 176 ANPLVVAEKVKV 187
DB 148 QNEPLVAKDKITK 159

RESULT 6
O7Y1W9_ORYSA PRELIMINARY; PRT; 161 AA.

AC O7Y1W9;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Putative zinc finger transcription factor ZFP33 (Putative protein 216).
DE Name:OSUNB0050P10.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaraloideae; Oryzaceae; Oryza.
NCBI_Taxid=39947;
RX
RN
RP NUCLEOTIDE SEQUENCE.
RA Huang J., Zhang H.-S.;
RN Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
[2]
RN
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsunoto T., Katayose Y.;
RN Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RT clone:OSUNB0050P10.";
DR EMBL; AY282740; AAP37480.1; -; mRNA.
DR EMBL; AP005840; BAD31780.1; -; genomic_DNA.
DR Gramene; O7Y1W9; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01428; Zf-A20; 1.
DR Pfam; PF01428; Zf-AN1; 1.
SQ SEQUENCE 161 AA; 17680 MW; 1443570BF15DD149 CRC64;

Query Match 39.1%; Score 384.5; DB 2; Length 161
Best Local Similarity 42.4%; Pred. No. 5.4e-24; Matches 81; Conservative 23; Mismatches 52; Indels 5;

OY 1 MATE---RVSOETTSQAPGPGVMCKNLGCFGSGQATMGLGCKCYRETVM 58
DB 1 MAQESWKKEAEETGHTPEAPILCVNCGFFSGSMTNNKSCYRDFV- 53
OY 59 QAAQATSATAAAVQPPAPVHETKLTCEVERTMIIVPHOSSSYQODLVTP 118
DB 54 -----APV-----VEKKAFTPPASSSKTPLEPAKI 89
OY 119 AA--PSRPEPNRCSCKRKGVLTFKRCRCGNLYCALHRSYDKHTCTYD 176
DB 90 AAQEPKPPSRKCLSCRRKQGLTFKRCRCGTFCSMRYADSHKCTFDYKQVGREQIAK 149
OY 177 NPVVAEKVKV 187
DB 150 NPVVAEKVKV 160

RESULT 7

O8H0X0_ARATH PRELIMINARY; PRT; 173 AA.
AC O8H0X0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Expressed protein.
DE Name=At1G51200;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_Taxid=3702;
RX
RN
RP NUCLEOTIDE SEQUENCE.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Mir, Palm C.J., Bowser L., Jones T., Bann J., Carninci P., Cne.

RA Cheuk R., Chung M.K., Hayaeshizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.,
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT001984; AAN71995.1; -, mRNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR002653; ZnF_A20.
 DR InterPro; IPR000058; ZnF_AN1.
 DR Pfam; PF01754; ZF-A20; 1.
 DR Pfam; PF01428; ZF-AN1; 1.
 DR SMART; SM00154; ZnF_AN1; 1.
 SQ SEQUENCE 173 AA; 18459 MW; 6A8F9ABA00C3DEBD CRC64;
 Query Match 39.0%; Score 384; DB 2; Length 173;
 Best Local Similarity 44.1%; Pred. No. 6,4e-24;
 Matches 78; Conservative 34; Mismatches 45; Indels 20; Gaps 5;
 QY 15 PEGPVMCKNLGCFPGSOATMGLCSKCYRETWQAKMTALAEQATQAAQTATTAAYQPP 74
 DB 12 PEGPKLCTNNCGFPFSAATNMCKCHKDMLFQ-----QEGGTFFASVSGTSSS----- 61
 QY 75 APVHETKLTCEVERTMIVPHQSSSYQODLVTPAAAPQAVKSSIAAP--SRPE-PNRCG 130
 DB 62 -----SNIKETFTALVDIETKSVPEMTVSVPSSVQVV-AEVVAPFEAAKPKGPSRCT 115
 QY 131 SCRKRVGLTGFKRCGNLYCALHRYSDKRTCTTYDYKKAQGEALIAANPLVNAEKYVK 187
 DB 116 TCKNRVGLTGFKRCGSLFCGTHRYADVHDCSFNYHAAQEAIAKAMPVKAELDK 172
 RESULT 8
 Q9SYC3_ARATH PRELIMINARY; PRT; 173 AA.
 ID Q9SYC3;
 AC Q9SYC3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
 DE FILM15.7 protein (At1g51200/FILM15_6) (At1g51200).
 OS Name=FILM15.7;
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid1; Brassicales; Brassicaceae; Arabidopsie.
 OC NCBI_TaxId=3702;
 RN NUCLEOTIDE SEQUENCE.
 RA Federhpiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Altafi H., Araujo R., Hultzer L., Rowley D., Buehler E., Dunn P.,
 RA Gonzalez A., Kremetsekata I., Kim C., Lenz C., Li J., Liu S.,
 RA Luros S., Schwartz J., Shim P., Toriumi M., Vysotskaya V.S.,
 RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bower L., Carinini P., Dale J.M., Goldsmith A.D., Hayaeshizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onders C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.,
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RA Cheuk R., Chen H., Kim C.J., Shim P., Ecker J.R.,
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC006085; AAD30634.1; -, Genomic DNA.
 DR EMBL; AY056445; AAL08301.1; -, mRNA.
 DR EMBL; BT015115; AAT71987.1; -, mRNA.
 DR PIR; G96549; G96549.

DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR002653; ZnF_A20.
 DR InterPro; IPR000058; ZnF_AN1.
 DR Pfam; PF01754; ZF-A20; 1.
 DR Pfam; PF01428; ZF-AN1; 1.
 DR SMART; SM00154; ZnF_AN1; 1.
 SQ SEQUENCE 173 AA; 18429 MW; 6A8F6CA09B3CBED CRC64;
 Query Match 38.5%; Score 379; DB 2; Length 173;
 Best Local Similarity 43.5%; Pred. No. 1.7e-23;
 Matches 77; Conservative 34; Mismatches 46; Indels 5;
 QY 15 PEGPVMCKNLGCFPGSOATMGLCSKCYRETWQAKMTALAEQATQAAQ 74
 DB 12 PEGPKLCTNNCGFPFSAATNMCKCHKDMLFQ-----QEGGAKFAP 61
 QY 75 APVHETKLTCEVERTMIVPHQSSSYQODLVTPAAAPQAVKSSIAAP-- 130
 DB 62 -----SNIKETFTALVDIETKSVPEMTVSVPSSVQVV-AEVVAPFE 115
 QY 131 SCRKRVGLTGFKRCGNLYCALHRYSDKRTCTTYDYKKAQGEALIAANPL 187
 DB 116 TCKNRVGLTGFKRCGSLFCGTHRYADVHDCSFNYHAAQEAIAKAMPV 172
 RESULT 9
 Q66IP3_XENLA PRELIMINARY; PRT; 211 AA.
 ID Q66IP3;
 AC Q66IP3;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE MGC86388 protein.
 GN Name=MGC86388;
 OS Xenopus laevis (African clawed frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipid
 OC Xenopodine; Xenopus; Xenopus.
 OC NCBI_TaxId=8355;
 RN NUCLEOTIDE SEQUENCE.
 RA Kleins S.L., Strausberg R.L., Wagner L., Pontius J., Clift
 RA Richardson P.,
 RT "Genetic and genomic tools for Xenopus research: The NIH X
 RT initiative."
 RT dev. Dyn. 225:384-391(2002).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RA TISSUE=Kidney;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603
 RX STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmer C.M., Schu
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bha
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsi
 RA Diachenko L., Marueta K., Farmer A.A., Rubin G.M., Hong
 RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant J.L., Sc
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinini P., Pr
 RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mu
 RA Boask S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunara
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard
 RA Rodriguez A.C., Grimwood J., Green E.D., Dickson M.C.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).


```

RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC081266; AAH81266.1; -; mRNA.
DR SMR; Q661P3; 138-194.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01754; zf-A20; 1.
DR Pfam; PF01428; zf-AN1; 1.
DR SMART; SM00259; Znf_A20; 1.
DR SMART; SM00154; Znf_AN1; 1.
SQ SEQUENCE 211 AA; 22887 MW; F08CBF1EC3789964 CRC64;

Query Match 37.5%; Score 369; DB 2; Length 211;
Best Local Similarity 37.9%; Pred. No. 1,4e-22;
Matches 81; Conservative 29; Mismatches 64; Indels 40; Gaps 7;

QY 6 VSQETTSQAPGPGVWCKNLGFFGSOATWGLGSKCYRETWQAKMTALAO-ATQAAQAT 64
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   1 MAOE-TNQT-P-GPMLCNTGCGFYGNPRTNGMCSVCYKHLQKNSGRISIPMGASGSNRP 58
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

DB 65 SATAAAVQPPAPVHETKTCTE-----VERTMIVPHOS 96
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 59 SAESASVQGV-----ETSLNCEGAVGGLSDKSRNTPLAALPTVQMTMSISREDHVASPK 114
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 97 SSVYQDQDLVT---PAAAPAPVAVKSSIAAPSRPEP--NRGSCGKRVGLTGFKRCGNLYCA 151
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 115 TETSEPVVTQSPSPSYAOSTSLNKEKEPELPKPKKRCMKCKKJGLTFDRCGNLFCG 174
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 152 LHRYSDKHTCTYDYKAGQGAIAKAPLVVAEKV 185
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 175 LHRYSDKHNCPPDYKAAAKIRKENPVVAEKI 208
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
Q6DF90_XENLA PRELIMINARY; PRT; 211 AA.
ID Q6DF90;
AC Q6DF90;
RT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
GN Name=za20d2-prov;
DE Name=za20d2-prov;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_TaxID=8335;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Oocytes;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

```

```

RT "Generation and initial analysis of more than 15,000 full-
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RA Strausberg R.L., Wagner L., Pontius J., Cliftc
RA Klein S.L., Scherch R.L., Wagner L., Pontius J., Cliftc
RT "Genetic and genomic tools for Xenopus research: The NIH ;
RT Initiative.";
RT Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Oocytes;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC076851; AAH76851.1; -; mRNA.
DR SMR; Q6DF90; 138-194.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01754; zf-A20; 1.
DR Pfam; PF01428; zf-AN1; 1.
DR SMART; SM00259; Znf_A20; 1.
DR SMART; SM00154; Znf_AN1; 1.
SQ SEQUENCE 211 AA; 22974 MW; A6FC08103F36DBD CRC64;

Query Match 37.1%; Score 365; DB 2; Length 211;
Best Local Similarity 36.9%; Pred. No. 2.9e-22;
Matches 79; Conservative 32; Mismatches 63; Indels 7;

QY 6 VSQETTSQAPGPGVWCKNLGFFGSOATWGLGSKCYRETWQAKMTAL.
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MAOE-TNQT-P-GPMLCNTGCGFYGNPRTNGMCSVCYKHLQKNSGR.
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 65 SATAAAVQPPAPVHETKTCTE-----VERTMIVPHO.
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

DB 59 STESASIQRV-----ETSLNCEGAAGLSDKSRNTPLAALPTVQMTM.
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 108 AAPAV-----KSSIAAPS-----RPENRGSCGKRVGLT.
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 115 TETSEPVVTQSPSPSYAOSTSLNKEKEPELPKPKKRCMKCKKJGLT.
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 152 LHRYSDKHTCTYDYKAGQGAIAKAPLVVAEKV 185
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 175 LHRYSDKHNCPPDYKAAAKIRKENPVVAEKI 208
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
220D2_MOUSE STANDARD; PRT; 213 AA.
ID 220D2_MOUSE
AC 088878;
RT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zinc finger A20 domain containing protein 2 (Zinc finger ;
GN Name=za20d2; Synonyms=zf216, Znf216;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; S.
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE. AND TISSUE SPECIFICITY.
RP MEDLINE=98428793; PubMed=9758550; DOI=10.1016/S0378-1119(
RA Scott D.A., Greinwald U.H. Jr., Marietta J.R., Drury S.,
RA Swiderki R.B., Vinas A., DeAngelis M.M., Carmi R., Ramesh
RA Kraft M.L., Elbedour K., Skworak A.B., Friedman R.A.,
RA Srikanth Sridharan C.R., Verhoeven K., Van Camp G.,
RA Delinger P.L., Batzer M.A., Morton C.C., Keats B.J., Smi;
RA Sheffield V.C.;

```


RA Rogers J., Dunham I.;
RT "DNA sequence and analysis of human chromosome 9.";
RN Nature 429:369-374(2004).
RN (3)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Eye, Lung, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: Expressed in fetal cochlea. Also expressed in
infant brain, fetal heart, lung, pancreatic islet, melanocyte,
pituitary gland, placenta, corneal stroma, and parathyroid tumor.
CC -1- SIMILARITY: Contains 1 A20-type zinc finger.
CC -1- SIMILARITY: Contains 1 AN1-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL, AF062072; AAC61801.1; -; Genomic_DNA.
DR EMBL, AF062346; AAC42601.1; -; mRNA.
DR EMBL, AF062347; AAC42602.1; -; mRNA.
DR EMBL, AL135924; CAD13440.1; -; Genomic_DNA.
DR EMBL, BC011018; AAH11018.1; -; mRNA.
DR EMBL, BC027077; AAH27707.1; -; mRNA.
DR EMBL, BC073131; AAH73131.1; -; mRNA.
DR SMR, O76080; 140-196.
DR Ensemble, ENSG00000107372; Homo sapiens.
DR HGNC, HGNC:13008; ZA20D2.
DR MIM, 604761; -;
DR InterPro, IPR002653; Znf_A20.
DR InterPro, IPR000058; Znf_AN1.
DR Pfam, PF01754; zf-A20; 1.
DR Pfam, PF01428; zf-AN1; 1.
DR SMART, SM00259; Znf_A20; 1.
DR SMART, SM00154; Znf_AN1; 1.
DR PROSITE, PS1036; ZF_A20; 1.
DR PROSITE, PS1039; ZF_AN1; 1.
DR Metal-binding; Zinc; Zinc-finger.
FT ZN FING 8 42 A20-type.
FT ZN FING 151 194 AN1-type.
SQ SEQUENCE 213 AA; 23132 MW; E477504B1BA77753 CRC64;
Query Match 37.0%; Score 364; DB 1; Length 213;
Best Local Similarity 38.2%; Pred. No. 3,6e-22;
Matches 81; Conservative 31; Mismatches 66; Indels 34; Gaps 7;
QY 6 VSOETTSQAPBGPVWCKNLGCPFGSQATMGCLSCYKRETVMAQKXTALAEOATQAQATS 65
DB 1 MNOE-TNQP-TPMCLSTGCGFYGNPRINGMCSVCYKELHQQOQNSGRMPTGASGNS 58
QY 66 AT-AAA-VO-----PPA-----PHERKTKLTCVEVTRTIVIPHOSS 98

DB 59 PRTSDASVORADTSLNCCGAAAGSTSEKSRNVPVAAALPYTQOTKMSI; 118
QY 99 YQODLVTV---PAAAAPOAVKSSIAPSRDEP--NRGSGCRKRVGLTGF; 153
DB 119 VSEPVVTPSPSVSGSTSGSEKAPELPKPKNRNCFMCRKKTGGLGFI; 178
QY 154 RYSDKRTCTYDKAQQEAIAPANPLVAEKV 185
DB 179 RYSDKRNCPDYKAAEAAAKIRKNPVAEKI 210
RESULT 13
Q6PBK7_BRARE
ID Q6PBK7_BRARE PRELIMINARY; PRT; 213 AA.
AC Q6PBK7;
DT 05-JUL-2004 (TRENBLREL, 27, Created)
DT 05-JUL-2004 (TRENBLREL, 27, Last sequence update)
DE 05-JUL-2004 (TRENBLREL, 27, Last annotation update)
DE Zinc finger, A20 domain containing 2.
GN Name=za20d2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele.
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyp.
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Wild-type; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.24260;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Sch.
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., S.
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., P.
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., M.
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunar.
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Wild-type; TISSUE=Eye;
RL Director MGC Project;
RL Submitted (OCT-2003) to the EMBL/Genbank/DDJ databases.
DR EMBL, BC059673; AAH59673.1; -; mRNA.
DR SMR, Q6PBK7; 140-196.
DR ZFIN, ZDB-GENE-040426-1979; za20d2.
DR GO, GO:0003677; F:DNA binding; IEA.
DR GO, GO:0008270; F:zinc ion binding; IEA.
DR InterPro, IPR002653; Znf_A20.
DR InterPro, IPR000058; Znf_AN1.
DR Pfam, PF01754; zf-A20; 1.
DR Pfam, PF01428; zf-AN1; 1.
DR SMART, SM00259; Znf_A20; 1.
DR SMART, SM00154; Znf_AN1; 1.
SQ SEQUENCE 213 AA; 23280 MW; DC5626C21BA56BA3 CRC64;
Query Match 37.0%; Score 364; DB 2; Length 213;
Best Local Similarity 38.1%; Pred. No. 3,6e-22;
Matches 82; Conservative 33; Mismatches 60; Indels 7;
QY 6 VSOETTSQAPBGPVWCKNLGCPFGSQATMGCLSCYKRETVMAQKXTALAEOATQAQATS 65
DB 1 MNOE-TNQP-TPMCLSTGCGFYGNPRINGMCSVCYKELHQQOQNSGRMPTGASGNS 58
QY 66 AT-AAA-VO-----PPA-----PHERKTKLTCVEVTRTIVIPHOSS 98

DB 1 MAOE-TNOSP-V-PILCTTGCGFYGNPRTNGMCSVCYKEHLNRQSSDRSPMSFLAGSPSA 58
 QY 60 AAQATSAATAA-VQ-PAPVHER-----KLTCEVERMTIVHOSSSYQ 101
 DB 59 BASAQRLBASINKAETLPAPSTETMRBSIPSSLPVQKMT---EMSLSREKALSPXA 115
 QY 102 DLVTAPAAAPQAVKSSIAA-----PSPPEPNCGSCRKRVGLTGFRCRCGNLYC 150
 DB 116 ETVREVITVQPTSSYPPIPAQAQANDAKSPDSSKPKKNCFTCRKXVGLTGFPCCGNLFC 175
 QY 151 ALHRYSDKHCTTYDYKAAQGAIAKANPLVAAKV 185
 DB 176 GHRYSDKHCTYDYKABAAAKIRKENPVVADKI 210

RESULT 14
 ID Q5ZJR4 CHICK PRELIMINARY; PRT; 212 AA.
 AC Q5ZJR4;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=RCJMB04.16e14;
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 NCBI_TaxID=9031;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CB; TISSUE=Bursa;
 RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zain J.,
 RA Fiedler J., Kutter S., Blagoderaki A., Kostovska D., Kotet M.,
 RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.W.,
 RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
 RT gene function analysis."
 RL Genome Biol. 6:R6-R6(2005).
 DR EMBL, AJ720370; CAG32029.1; -; mRNA.
 DR SMR, Q5ZJR4; 139-195.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR002653; Znf_A20.
 DR InterPro; IPR000058; Znf_AN1.
 DR Pfam; PF01428; ZF-A20; 1.
 DR Pfam; PF01428; ZF-A20; 1.
 DR SMART; SM00259; Znf_A20; 1.
 DR SMART; SM00154; Znf_AN1; 1.
 DR SMART; SM00154; Znf_AN1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 212 AA; 23141 MW; D974C13CF9B935A9 CRC64;

Query Match 36.9%; Score 363.5; DB 2; Length 212;
 Best Local Similarity 38.0%; Pred. No. 3.9e-22;
 Matches 84; Conservative 27; Mismatches 57; Indels 53; Gaps 8;

QY 6 VSQETTSQAPBGPVWCKNLGFPFGSOATWGLCSKCYR----- 43
 DB 1 MTQE-TNCTP-GPMICSTGCGFYGNPRTNGMCSVCYKEHLNRQSSDRSPMSFLAGSPSA 58
 QY 44 -----TWQAKMTALAEQATQAQATSAATAA-VQ-PV-----HETKLTCEV 86
 DB 59 PTSBSTSVORADYTLN-NNCDGAVGSTSEKRSVPVAAALPVQOMTEMSISREKVTPT 117
 QY 87 ERTMTVPHOSSSYQODLVTPAAAPQAVKSSIAAPSRPEP--NRGSCRKRVGLTGFPCR 144
 DB 118 ETEPVVTPP-----TPSVSQPSTSRNEKAPELPKPKKRCFMCRCXVGLTGFDOR 168
 QY 145 CGNLFCALHRYSDKHCTTYDYKAAQGAIAKANPLVAAKV 185
 DB 169 CGNLFCALHRYSDKHCTTYDYKAAQGAIAKANPLVAAKV 209

Result 15

Q9LX15 ARATH
 ID Q9LX15 ARATH PRELIMINARY; PRT; 170 AA.
 AC Q9LX15;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Zinc finger-1like protein.
 GN Name=FCZ22.200;
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trach
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Purnelle B., Masny D., Goffeau A., Boutry M., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.,
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL353912; CAB89241.1; -; Genomic DNA.
 DR PIR; T49033; T49033.
 DR SMR; Q9LX15; 107-151.
 DR InterPro; IPR002653; Znf_A20.
 DR InterPro; IPR000058; Znf_AN1.
 DR Pfam; PF01428; ZF-A20; 1.
 DR SMART; SM00154; Znf_AN1; 1.
 SQ SEQUENCE 170 AA; 18598 MW; 81B2829C8F83D620 CRC64;

Query Match 36.8%; Score 362.5; DB 2; Length 212;
 Best Local Similarity 37.8%; Pred. No. 3.8e-22;
 Matches 73; Conservative 32; Mismatches 53; Indels 4;

QY 6 VSQETTSQAPBGPVWCKNLGFPFGSOATWGLCSKCYRRTVQAKMTALA 65
 DB 1 MAEHRQCTPESNRLCVNNGCFLGSSATMNLCSNCGDLCXK-----C 54
 QY 66 ATAAAVQPPAVVHERKLTCEVERMTIVHOSSSYQODLVTPAAAPQAV 119
 DB 55 ESSISVSPPS-----SSSSSEISSPILPPLKNPSVK 96
 QY 120 APS-----RPEPNCGSCRKRVGLTGFRCRCGNLYCALHRYSDKHCTTY 174
 DB 97 LPTBQNOQQRPNKCTCRKXVGLTGFRCRCGNLYCALHRYSDKHCTTY 156
 QY 175 KANPLVAAKVXK 187
 DB 157 KANPLVAAKXK 169

Search completed: December 8, 2005, 16:32:01
 Job time : 162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 15:03:00 ; Search time 44 Seconds
(without alignments)
353.251 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATERVSOETTSGAPEGVPM.....GGEAIKANPLVAEKVKVF 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2_6/prodata/1/1aa/5_COMB.pep:*
 - 2: /cgn2_6/prodata/1/1aa/6_COMB.pep:*
 - 3: /cgn2_6/prodata/1/1aa/H_COMB.pep:*
 - 4: /cgn2_6/prodata/1/1aa/ECTUS_COMB.pep:*
 - 5: /cgn2_6/prodata/1/1aa/RB_COMB.pep:*
 - 6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984	100.0	188	2	US-09-828-303-20 Sequence 20, Appl
2	366	37.2	213	2	US-09-244-805-11 Sequence 11, Appl
3	365	37.1	395	2	US-09-949-016-9564 Sequence 9564, Ap
4	364	37.0	213	2	US-09-949-016-6681 Sequence 6681, Ap
5	343.5	34.9	208	1	US-08-861-269-3 Sequence 3, Appl1
6	343.5	34.9	208	1	US-09-134-596-3 Sequence 3, Appl1
7	343.5	34.9	208	2	US-09-293-273-3 Sequence 3, Appl1
8	313	31.8	189	1	US-08-861-269-7 Sequence 7, Appl1
9	313	31.8	189	1	US-09-134-596-7 Sequence 7, Appl1
10	313	31.8	189	1	US-09-293-273-7 Sequence 7, Appl1
11	177	18.0	46	1	US-08-861-269-8 Sequence 8, Appl1
12	177	18.0	46	1	US-09-134-596-8 Sequence 8, Appl1
13	177	18.0	46	2	US-09-293-273-8 Sequence 8, Appl1
14	101	10.3	497	1	US-09-134-000C-5990 Sequence 5990, Ap
15	99.5	10.1	491	2	US-08-942-819-2 Sequence 2, Appl1
16	99.5	10.1	491	2	US-09-522-955A-2 Sequence 2, Appl1
17	98	10.0	87	2	US-09-513-999C-5919 Sequence 5919, Ap
18	98	10.0	258	2	US-10-012-231A-153 Sequence 153, App
19	98	10.0	258	2	US-10-015-389A-153 Sequence 153, App
20	98	10.0	258	2	US-10-006-768A-153 Sequence 153, App
21	98	10.0	258	2	US-10-015-671A-153 Sequence 153, App
22	98	10.0	258	2	US-10-015-671A-153 Sequence 153, App
23	98	10.0	258	2	US-10-011-833A-153 Sequence 153, App
24	98	10.0	258	2	US-10-006-041A-153 Sequence 153, App
25	98	10.0	258	2	US-10-012-064A-153 Sequence 153, App
26	93.5	9.4	480	2	US-09-108-020-6 Sequence 6, Appl1
27	93.5	9.4	480	2	US-09-685-296-6 Sequence 6, Appl1

28	92	9.3	414	2	US-09-270-767-57349	Sec	' A
29	92	9.3	506	2	US-09-270-767-42086	Sec	' A
30	89	9.0	174	2	US-09-248-796A-25211	Sec	' A
31	88	8.9	595	1	US-08-232-087A-2	Sec	' A
32	88	8.9	595	2	US-09-006-353A-9	Sec	' A
33	88	8.9	595	2	US-09-573-986-9	Sec	' A
34	88	8.9	595	2	US-09-949-016-6048	Sec	' A
35	88	8.9	642	2	US-09-949-016-8043	Sec	' A
36	87.5	8.9	176	2	US-10-104-047-2419	Sec	' A
37	87	8.8	595	1	US-08-570-923-2	Sec	' A
38	87	8.8	595	1	US-08-225-989-2	Sec	' A
39	87	8.8	595	1	US-08-580-014-2	Sec	' A
40	87	8.8	595	2	US-09-079-785-2	Sec	' A
41	87	8.8	595	2	US-09-921-667-6	Sec	' A
42	86.5	8.8	595	2	US-09-628-126-2	Sec	' A
43	86.5	8.8	361	2	US-09-976-594-351	Sec	' A
44	85.5	8.7	456	2	US-09-689-486-56	Sec	' A
45	84.5	8.6	792	2	US-09-949-016-10798	Sec	' A

ALIGNMENTS

```
RESULT 1
US-09-828-303-20
; Sequence 20, Application US/09828303
; Patent No. 6677504
; GENERAL INFORMATION:
; APPLICANT: COSTA R SILVA, OSWALDO DA
; APPLICANT: BOHNETT, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PRO.
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-303-20

Query Match          100.0%; Score 984; DB 2; Length 1
Best Local Similarity 100.0%; Pred. No. 1e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0;

QY      1 MATERVSOETTSGAPEGVPMCKNLGFGFSQATMGICSKCYRETVMOA
DB      1 MATERVSOETTSGAPEGVPMCKNLGFGFSQATMGICSKCYRETVMOA
QY      61 AQAISATAAVQPPAPVHETLTCEVERTWIVPHOSSSYOODLVTPAAV
DB      61 AQAISATAAVQPPAPVHETLTCEVERTWIVPHOSSSYOODLVTPAAV
QY      121 PSRPPNRCGSCRKRVGLTGFKRCGNLYCALHYSDDHGTCTYKAC
DB      121 PSRPPNRCGSCRKRVGLTGFKRCGNLYCALHYSDDHGTCTYKAC
QY      181 VAEKVKVF 188
DB      181 VAEKVKVF 188

RESULT 2
US-09-244-805-11
; Sequence 11, Application US/09244805
; Patent No. 6699660
; GENERAL INFORMATION:
```

```

? APPLICANT: Worley, Paul F.
? APPLICANT: Ianahan, Anthony
? APPLICANT: Goetz, Bernard
? APPLICANT: Helmisch, Holger
? APPLICANT: Kuner, Rohini
? APPLICANT: Scheek, Sigrid
? APPLICANT: Nikolic, Karoly
? APPLICANT: Zhukovski, Eugene
? TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE
? THEREFOR
? FILE REFERENCE: 10496/004001
? CURRENT APPLICATION NUMBER: US/09/244,805
? CURRENT FILING DATE: 1999-02-05
? PRIOR APPLICATION NUMBER: 60/074,518
? PRIOR FILING DATE: 1998-02-12
? PRIOR APPLICATION NUMBER: 60/074,135
? PRIOR FILING DATE: 1998-02-06
? NUMBER OF SEQ ID NOS: 62
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 11
? LENGTH: 213
? TYPE: PRT
? ORGANISM: Eukaryote
?-09-244-805-11

```

[illegible]

```

RESULT 3
US-09-949-016-9564
/ Sequence 9564, Application US/09949016
/ Patent No. 681239
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9564
/ LENGTH: 395
/ TYPE: PRT
/ ORGANISM: Human
/ US-09-949-016-9564

```

Query Match	37.1%	Score 365;	DB 2;	Length 395;
Best Local Similarity	37.9%	Pred. No. 8.1e-31;		

[illegible]

```

RESULT 4
US-09-949-016-6681
; Sequence 6681, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6681
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6681

```

Query Match	Similarity	37.0%	Score 364	DB 2	Length 213
Best Local	Similarity	38.2%	Pred. No. 4.56-31		
Matches	Conservative	31	Mismatches	66	Indels
OY	6	VSQETTSQAPGCPVWCNNLCGFFSGSQATWGLGSCKCYRETNQWQAKTALA			65
Db	1	MAQD-TNQDTP-GPMLCTGCGCFYGNPRTNMGCSVCYEKHLQROQNSGRM			58
OY	66	AT--AAVQ-----PPA--PVHETKLTCEVE			98
Db	59	PTSDSASVORADTSLNCCGAAGSTSEKSRNVPAALPVYQOMTMSIS			118
OY	99	YQODLVT---PAAAPQAVKSSIAAPSRPEP--RRGSCSRKRVLLTGFK			153
Db	119	VSEPVVTQPSBSVSQPSSTQSEEAQPELPPKPKKRCRCMKCKKQVLTGFD			178
OY	154	RYSDKHTCTTYDYKAAGQBAIAKANPLVVAEKV 185			
Db	179	RYSDKHNCPYDYKAAAKIRKKNPVVAAEKI 210			

RESULTS
US-08-861-269-3
; Sequence 3, Application US/08061265
; Patent No. 5817494
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti

APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,269
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 208 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-08-861-269-3

Query Match 34.9%; Score 343.5; DB 1; Length 208;
Best Local Similarity 36.4%; Pred. No. 7e-29;
Matches 78; Conservative 31; Mismatches 66; Indels 39; Gaps 7;
QY 6 VSGEFTSQAPGEGPVWCKNLGFGSGQATMGLCKCYR-----TYMAK 49
DB 1 MAQETNHS--QVPMCLSTGCGFYGNPRTNGMCSVCYKEHLQRNNGSNGR 58
QY 50 MTAALAEQAT-----QAQATSTATAAVQPPAPVHETKLTCEVERITMIVPHOSSVQODLV 104
DB 59 SESLPVQCTDGSVPFAQSLDSTSSMO--PSPVSNQSLSE---SVASSQDSTSDVKA 113
QY 105 TPAALAPQA-VKSSIAAPS-----RPEPNRCSCKRKVGTLGFKRCGNLYCALH 153
DB 114 VPETEDVQASVSDTAQGPSEBQSKLEKPKQKRCFMCKRKVGLGFGRCGNLYCGVH 173
QY 154 RYSDKHTCTTYDYKAAQGEAIAKANPLVVAEKVVK 187
DB 174 RYSDVHNCSSYNYKADAERIKRENPPVVGEEKIOK 207

RESULT 6
US-09-134-596-3
Sequence 3, Application US/09134596
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,596
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/861,269
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 208 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-134-596-3

Query Match 34.9%; Score 343.5; DB 1; Length 208;
Best Local Similarity 36.4%; Pred. No. 7e-29;
Matches 78; Conservative 31; Mismatches 66; Indels 39;
QY 6 VSGEFTSQAPGEGPVWCKNLGFGSGQATMGLCKCYR-----TYMAK 49
DB 1 MAQETNHS--QVPMCLSTGCGFYGNPRTNGMCSVCYKEHLQRNNGSNGR 58
QY 50 MTAALAEQAT-----QAQATSTATAAVQPPAPVHETKLTCEVERITMIV 104
DB 59 SESLPVQCTDGSVPFAQSLDSTSSMO--PSPVSNQSLSE---SVASSQDSTSDVKA 113
QY 105 TPAALAPQA-VKSSIAAPS-----RPEPNRCSCKRKVGTLGFKRCGNLYCALH 153
DB 114 VPETEDVQASVSDTAQGPSEBQSKLEKPKQKRCFMCKRKVGLGFGRCGNLYCGVH 173
QY 154 RYSDKHTCTTYDYKAAQGEAIAKANPLVVAEKVVK 187
DB 174 RYSDVHNCSSYNYKADAERIKRENPPVVGEEKIOK 207

RESULT 7
US-09-293-273-3
Sequence 3, Application US/09293273
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/293,273
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/134,596
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT09
; CLONE: 2151473
;
US-09-293-273-3

```

```

Query Match          34.9%; Score 343.5; DB 2; Length 208;
Best Local Similarity 36.4%; Pred. No. 7e-29;
Matches 78; Conservative 31; Mismatches 66; Indels 39; Gaps 7;

QY 6 VSQETSAPEGPVCKNLGFFGSOATWGLCSKYRE-----TWQAK 49
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 1 MAQETNHS--QVPMLCSTCGFGYGNPRTGMCVYKEHLQRONSNGRISPPATSVSSL 58
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 50 MTALEAQT-----QAAQTSATTAAYOPAPVHETKLTCEERTMTIVPHOSSVQQDLV 104
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 59 SSSLVQCTDGSVPEHQSLDSTSSMQ--PSFVSNOSLSL---SVASSQLDSTSDKA 113
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 105 TPAAPAPQAVKSIAPB-----RPPNRCGSCRKRVGLTGFKRCGNLYCALH 153
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 114 VPEDEDVQASVSDTAQQPSEBQSKLEKPKQKQKRCFMKRVGLTGFECCRGVYCGVH 173
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 154 RYSDKHTCTYDKAAGEIAKAPLYVAEKYVK 187
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 174 RYSDVHNCSTYNYKADAEXIKRENPNVVGKIQK 207
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

```

```

RESULT 8
US-08-861-269-7
; Sequence 7, Application US/08861269
; Patent No. 5817494
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIOUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,269
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1279278
;
US-08-861-269-7

```

```

Query Match          31.8%; Score 313; DB 1; Length 189
Best Local Similarity 35.4%; Pred. No. 1.2e-25;
Matches 68; Conservative 31; Mismatches 69; Indels 6;

QY 8 QETTSQAPEGPVCKNLGFFGSOATWGLCSKYRETYWQAKMTALAEQ 67
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 5 QOQQTAPS-----CRAGGFPGASATGEGYCSQCFKNTLKRQODVRLTS 60
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 68 AAAYQ-PPAPV-----HETKLTCEERTMTIVPHOSSVQQDL 113
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 61 SSALKSEPSYDMCMKAAYSDETAQVDC-----DIINVCDOINDS 115
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 114 VKSIIAPSRPEPNRCGSCRKRVGLTGFKRCGNLYCALHRYSDKHTCT 173
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 116 ITVDVPVPVK-KANRCHMKKRVGLTGFSCHGGLYGCSDHRYDDAHNC 174
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 174 AKANPLVVAERY 185
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 175 RKNPNVVVSDKV 186
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

```

```

RESULT 9
US-09-134-596-7
; Sequence 7, Application US/09134596
; Patent No. 5922318
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIOUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,596
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/861,269
;

```



```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1279278
; US-09-134-596-7

Query Match          31.8%; Score 313; DB 1; Length 189;
Best Local Similarity 35.4%; Pred. No. 1.2e-25;
Matches 68; Conservative 31; Mismatches 69; Indels 24; Gaps 6;

QY 8 QETSOAPEGPVMCKNLGCFPGSOATMGLCSKCYRETVMQAKMTALAEQATQAQTSAT 67
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5 QQAQTAAPS-----CRAGCGFPGASATBEGYCSQCFKNTLKRQDITRLTSPVSPSSMAAT 60
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 68 AAAPV-PPAPV-----HETKLTCEVERTMIVPHOSSVYQODLVTPA-AAAPQA 113
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 61 SSALSPSSVDMCKMAVSVSDETAAMDCE-----DIINVDDQINDBSVTVAESTAPTT 115
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 114 VKSSIAAPSRPEPNRCGSKRRVGLTGFCRCGNLYCALHRSYDHTCTYYDKAAQGEAI 173
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 116 ITVDVPVPEVK-KANRCHMCKRRVGLTGFSRCGGLYCGDHRVDOAHNCQFDYKTMERETI 174
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 174 AKANPLVVAEKV 185
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 175 KKNNPVVVSDKV 186
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 10
US-09-293-273-7
; Sequence 7, Application US/09293273
; Patent No. 6057112
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/293,273
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/134,596
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1279278
; US-09-293-273-7

Query Match          31.8%; Score 313; DB 2; Length 189;
Best Local Similarity 35.4%; Pred. No. 1.2e-25;
Matches 68; Conservative 31; Mismatches 69; Indels 24; Gaps 6;

QY 8 QETSOAPEGPVMCKNLGCFPGSOATMGLCSKCYRETVMQAKMTALAE 67
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5 QQAQTAAPS-----CRAGCGFPGASATBEGYCSQCFKNTLKRQDITRLT 60
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 68 AAAPV-PPAPV-----HETKLTCEVERTMIVPHOSSVYQODLVTPA-AAAPQA 113
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 61 SSALSPSSVDMCKMAVSVSDETAAMDCE-----DIINVDDQINDBSVTVAESTAPTT 115
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 114 VKSSIAAPSRPEPNRCGSKRRVGLTGFCRCGNLYCALHRSYDHTCTYYDKAAQGEAI 173
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 116 ITVDVPVPEVK-KANRCHMCKRRVGLTGFSRCGGLYCGDHRVDOAHNCQFDYKTMERETI 174
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 174 AKANPLVVAEKV 185
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 175 KKNNPVVVSDKV 186
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 11
US-08-861-269-8
; Sequence 8, Application US/08861269
; Patent No. 5817494
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,269
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
```

```
/ LENGTH: 46 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 998680
US-08-861-269-8

Query Match 18.0%; Score 177; DB 1; Length 46;
Best Local Similarity 69.8%; Pred. No. 7e-12;
Matches 30; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 143 CRCGNLYCALHRYSDKHTCTYYKAGGEAIAKAPLVVAEKV 185
DB 1 CRCGNLYCALHRYSDKHTCTYYKAGGEAIAKAPLVVAEKI 43

RESULT 12
US-09-134-596-8
/ Sequence 8, Application US/09134596
/ Patent No. 592318
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Corley, Neil C.
/ APPLICANT: Lal, Preeti
/ APPLICANT: Shah, Purvi
/ TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/134,596
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/861,269
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0302 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 46 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 998680
US-09-134-596-8

Query Match 18.0%; Score 177; DB 1; Length 46;
Best Local Similarity 69.8%; Pred. No. 7e-12;
Matches 30; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 143 CRCGNLYCALHRYSDKHTCTYYKAGGEAIAKAPLVVAEKV 185
DB 1 CRCGNLYCALHRYSDKHTCTYYKAGGEAIAKAPLVVAEKI 43
```

```
RESULT 13
US-09-293-273-8
/ Sequence 8, Application US/09293273
/ Patent No. 6057112
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Corley, Neil C.
/ APPLICANT: Lal, Preeti
/ APPLICANT: Shah, Purvi
/ TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/293,273
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/134,596
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0302 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 46 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 998680
US-09-293-273-8

Query Match 18.0%; Score 177; DB 2; Length 46;
Best Local Similarity 69.8%; Pred. No. 7e-12;
Matches 30; Conservative 3; Mismatches 10; Indels 0;

QY 143 CRCGNLYCALHRYSDKHTCTYYKAGGEAIAKAPLVVAEKV 185
DB 1 CRCGNLYCALHRYSDKHTCTYYKAGGEAIAKAPLVVAEKI 43

RESULT 14
US-09-134-000C-5990
/ Sequence 5990, Application US/09134000C
/ Patent No. 6617156
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RE
/ FILE REFERENCE: 032796-032
/ CURRENT APPLICATION NUMBER: US/09/134,000C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/055,778
/ PRIOR FILING DATE: 1997-08-15
/ NUMBER OF SEQ ID NOS: 6812
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5990
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 16:26:56 ; Search time 116 Seconds
(without alignments)
677.172 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSOETTSGAPEGVPM.....GGEAIKANPLVAKVKYKF 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBSCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBSCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBSCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBSCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBSCOMB.pep.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984	100.0	188	3	US-09-828-303-20 Sequence 20, Appl
2	984	100.0	188	4	US-10-716-089-20 Sequence 42138, A
3	425.5	43.2	171	4	US-10-767-701-47138 Sequence 123596, A
4	419.5	42.6	171	4	US-10-437-963-123596 Sequence 140292, A
5	416.5	42.3	173	4	US-10-437-963-140292 Sequence 140291, A
6	416.5	42.3	338	4	US-10-437-963-140292 Sequence 261492, A
7	415	42.2	193	4	US-10-425-115-261492 Sequence 261492, A
8	415	42.2	211	4	US-10-425-115-261492 Sequence 261494, A
9	415	42.2	224	4	US-10-425-115-261494 Sequence 266094, A
10	414.5	42.1	171	4	US-10-425-115-256094 Sequence 59729, A
11	414.5	41.9	179	4	US-10-425-114-59729 Sequence 65423, A
12	412.5	41.9	179	4	US-10-425-114-65423 Sequence 45710, A
13	411.5	41.8	171	4	US-10-767-701-45710 Sequence 30832, A
14	407.5	41.4	171	4	US-10-437-963-123595 Sequence 30832, A
15	406.5	41.3	174	4	US-10-425-115-308732 Sequence 256116, A
16	403	41.0	205	4	US-10-425-115-256116 Sequence 44878, A
17	402.5	40.9	173	4	US-10-425-114-44878 Sequence 57001, A
18	402.5	40.9	173	4	US-10-425-114-57001 Sequence 185981, A
19	401.5	40.8	171	4	US-10-425-115-185981 Sequence 185983, A
20	401.5	40.8	171	4	US-10-425-115-185983 Sequence 178876, A
21	401.5	40.8	172	4	US-10-424-599-178876 Sequence 57676, A
22	401.5	40.8	173	4	US-10-425-114-57676 Sequence 231901, A
23	395.5	40.2	164	4	US-10-424-599-231901 Sequence 270156, A
24	392.5	39.9	159	4	US-10-425-115-270156 Sequence 231898, A
25	391.5	39.8	160	4	US-10-424-599-231898 Sequence 231899, A
26	391.5	39.8	160	4	US-10-424-599-231899 Sequence 266773, A
27	390	39.6	247	4	US-10-425-115-266773 Sequence 266773, A

28	389	39.5	174	4	US-10-424-599-204286 Seq
29	386	39.2	160	4	US-10-437-963-154553 Seq
30	386	39.2	172	4	US-10-767-701-44663 Seq
31	384.5	39.1	161	4	US-10-437-963-112566 Seq
32	383.5	39.0	175	4	US-10-424-599-204285 Seq
33	374.5	38.1	172	4	US-10-424-599-203681 Seq
34	374	38.0	172	4	US-10-424-599-242209 Seq
35	366	37.2	213	3	US-09-244-805-11 Seq
36	366	37.2	213	3	US-09-245-277-11 Seq
37	366	37.2	213	3	US-10-792-481-11 Seq
38	365.5	37.1	172	4	US-10-424-599-242210 Seq
39	365	37.1	229	4	US-10-264-049-2758 Seq
40	365	37.1	310	3	US-09-925-300-1344 Seq
41	364	37.0	213	5	US-10-507-617-27 Seq
42	361	36.7	224	4	US-10-437-963-116025 Seq
43	357	36.3	163	4	US-10-425-115-369121 Seq
44	357	36.3	224	4	US-10-425-115-71695 Seq
45	355.5	36.1	198	4	US-10-425-115-256050 Seq

ALIGNMENTS

```

RESULT 1
US-09-828-303-20
; Sequence 20, Application US/09828303
; Patent No. US20020102695A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNETT, HANS J.
; APPLICANT: VAN THIELEN, NOCHIA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PRO.
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-303-20

Query Match      100.0%  Score 984;  DB 3;  Length 18
Best Local Similarity 100.0%;  Pred. NO. 6.7e-85;
Matches 188;  Conservative 0;  Mismatches 0;  Indels 0;

QY      1 MATEVSOETTSGAPEGVPMCKNLGFGSQATMGLCSKCYRETMQAK
      |||||
DB      1 MATEVSOETTSGAPEGVPMCKNLGFGSQATMGLCSKCYRETMQAK
      |||||
QY      61 AQAISATAAQAQPAVHEHTKLTCEVERTWIVPHOSSYQODLVTPAA
      |||||
DB      61 AQAISATAAQAQPAVHEHTKLTCEVERTWIVPHOSSYQODLVTPAA
      |||||
QY      121 PSRPPNRCGSCRKVGLTGFCRCGNLYCALHRSDDHTCTYDYKAK
      |||||
DB      121 PSRPPNRCGSCRKVGLTGFCRCGNLYCALHRSDDHTCTYDYKAK
      |||||
QY      181 VAEKVKYKF 188
      |||||
DB      181 VAEKVKYKF 188
      |||||

RESULT 2
US-10-716-089-20
; Sequence 20, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:

```

APPLICANT: COSTA E SILVA, OSWALDO DA
APPLICANT: BOHNET, HANS J.
APPLICANT: VAN THIELEN, NOCHA
APPLICANT: CHEN, ROUYING
TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
TITLE OF INVENTION: METHODS OF USE IN PLANTS
FILE REFERENCE: 16313-0030
CURRENT APPLICATION NUMBER: US/10/716,089
CURRENT FILING DATE: 2003-11-18
PRIOR APPLICATION NUMBER: 60/196,001
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 188
TYPE: PRT
ORGANISM: Physcomitrella patens
US-10-716-089-20

Query Match 100.0%; Score 984; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 6.7e-85;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATERVSGETTSQAEQPGVWMCNLCGFGSQATMGLCSKCYRETYWQAKMTALAEQATQA 60
DB 1 MATERVSGETTSQAEQPGVWMCNLCGFGSQATMGLCSKCYRETYWQAKMTALAEQATQA 60
QY 61 AQTATATAAVPPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
DB 61 AQTATATAAVPPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
QY 121 PSRPPNCGSRKRVGLTFKRCRCGNLYCALHRYSDKHCTYDYKAAQGEAIKAMPV 180
DB 121 PSRPPNCGSRKRVGLTFKRCRCGNLYCALHRYSDKHCTYDYKAAQGEAIKAMPV 180
QY 181 VAEKVKVF 188
DB 181 VAEKVKVF 188

RESULT 3
US-10-767-701-47138
Sequence 47138, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 47138
LENGTH: 171
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28WAY03-C3087_1.pep
US-10-767-701-47138

Query Match 43.2%; Score 425.5; DB 4; Length 171;
Best Local Similarity 46.1%; Pred. No. 3.9e-32;
Matches 83; Conservative 28; Mismatches 56; Indels 13; Gaps 2;
QY 8 QETTSQAEQPGVWMCNLCGFGSQATMGLCSKCYRETYWQAKMTALAEQATQAQATSAT 67
DB 4 KETGQGPBGPILCINNGCFPSAATMMCKCHKEMIMKQAKLA--ASSIDSIVNGN 61
QY 68 AAAYOPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAAPSRPEPN 127
DB 62 DAVMEP-----VAGNTVVAADIEIQTMNVQPADVAGPSEGAALVSKGVGN 110

QY 128 RCGSRKRVGLTFKRCRCGNLYCALHRYSDKHCTYDYKAAQGEAIKAA 187
DB 111 RCGSRKRVGLTFKRCRCGNLYCALHRYSDKHCTYDYKAAQGEAIKAA 170

RESULT 4
US-10-437-963-123596
Sequence 123596, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Mc
TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 123596
LENGTH: 171
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRTA530_26415C.1.pep
US-10-437-963-123596

Query Match 42.6%; Score 419.5; DB 4; Length 1
Best Local Similarity 45.6%; Pred. No. 1.5e-31;
Matches 82; Conservative 26; Mismatches 59; Indels 2;

QY 8 QETTSQAEQPGVWMCNLCGFGSQATMGLCSKCYRETYWQAKMTALAEQ 67
DB 4 KETGQGPBGPILCINNGCFPSAATMMCKCHKEMIMKQAKLA-- 61
QY 68 AAAYOPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVK 127
DB 62 DSGKEPIIAGHAEVAAQVEVKTIV-----AQPAEIAQPSG 110
QY 128 RCGSRKRVGLTFKRCRCGNLYCALHRYSDKHCTYDYKAAQGEAIKAA 187
DB 111 RCGSRKRVGLTFKRCRCGNLYCAMHRYSDKHCTYDYKAAQGEAIKAA 170

RESULT 5
US-10-437-963-140292
Sequence 140292, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Mc
TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 140292
LENGTH: 173
TYPE: PRT
ORGANISM: Oryza sativa


```
Db      41 KEAGCAEBGPILCINNCGFFGSAATNMCKCHKEMITTKQOAKLA--ASSIDSVNGS 98
Qy      68 AAAVQPPAPVHETKLTCEVERTMIIVPHOSSSYQODLVTPA-AAAPQAVKSSIAAPSRPEP 126
      99 DAVMEPPVAVGNT-----VVAVAQVELQTMNVQO---PADVAGSSEGVAAISKGGKVG 149
Db      127 NRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAKAGOBIAKANPLVVAEKV 186
      150 NRCSACRKRKRVGLTGFKRCGNLYCALHRYSDKHDCKFDYRTAARDALIAKANPVVADKLD 209
Qy      187 K 187
      210 K 210
Db

RESULT 9
US-10-425-115-261494
; Sequence 261494, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261494
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170097C.1.pcp
US-10-425-115-261494

Query Match
Best Local Similarity 47.0%; Score 415; DB 4; Length 224;
Matches 85; Conservative 27; Mismatches 57; Indels 12; Gaps 4;

Qy      8 QETTSQAEPEGVPWCKNLGCFGSGQATMGLCSKCYRETVMAQMTALAEQATQAAQATCAT 67
      39 KEAGCAEBGPILCINNCGFFGSAATNMCKCHKEMITTKQOAKLA--ASSIDSVNGS 96
Db      68 AAAVQPPAPVHETKLTCEVERTMIIVPHOSSSYQODLVTPA-AAAPQAVKSSIAAPSRPEP 126
      97 DAVMEPPVAVGNT-----VVAVAQVELQTMNVQO---PADVAGSSEGVAAISKGGKVG 147
Db      127 NRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAKAGOBIAKANPLVVAEKV 186
      148 NRCSACRKRKRVGLTGFKRCGNLYCALHRYSDKHDCKFDYRTAARDALIAKANPVVADKLD 207
Qy      187 K 187
      208 K 208
Db

RESULT 10
US-10-425-115-256094
; Sequence 256094, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 256094
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(171)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_16514C.1.pcp
US-10-425-115-256094

Query Match
Best Local Similarity 43.1%; Score 414.5; DB 4; Length 171
Matches 84; Conservative 26; Mismatches 42; Indels 4;

Qy      8 QETTSQAEPEGVPWCKNLGCFGSGQATMGLCSKCYRETVMAQMTALAEBC 67
      4 KEAGCAEBGPILCINNCGFFGSAATNMCKCHKEMIMKQEOAKLA--
Db      68 AAAVQPPAPVHETKLTCEVERTMIIVPHOSSSYQODLVTPA-AAAPQAVK- 121
      62 DAVMEPPVAVGN-----TVVAADIELO- 95
Qy      122 S-----RPEPNRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTC 172
      96 SXGGGGLQREGXPNRCSGTCKRKRVGLTGFKRCGNLYCALHRYSDKHDCK 155
Db      173 IAKANPLVVAEKV 187
      156 IAKANPVVAKDKLCK 170
Db

RESULT 11
US-10-425-114-59729
; Sequence 59729, Application US/10425114
; Publication No. US2004003488BA1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; FILE REFERENCE: 38-21(53133)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59729
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3912-048-B5_FLI.pcp
US-10-425-114-59729

Query Match
Best Local Similarity 41.9%; Score 412.5; DB 4; Length 171
Matches 80; Conservative 28; Mismatches 50; Indels 3;

Qy      8 QETTSQAEPEGVPWCKNLGCFGSGQATMGLCSKCYRETVMAQMTALAEBC 60
      12 KEAGCAEBGPILCINNCGFFGSAATNMCKCHKEMIMKQEOAQLAAS 71
Db      61 --AQTSAATAAVQPPAPVHETKLTCEVERTMIIVPHOSSSYQODLVTPA 118
      72 GKGRALIAITGVAVP-----QVEKTI-----VQPM 109
Qy      119 AAPSRPEPNRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKA 178
```


Db 110 IAKAKEGPNRCATCRKRVGLTGFCNRCGNTYCSMHRYSDKHDCQPDYRTAARDALIAKANP 169

Qy 179 LVVAERVK 187

Db 170 VVKAERLKD 178

RESULT 12

US-10-425-114-65423

Sequence 65423, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 65423

LENGTH: 179

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: LIB4764-003-Fl1_Fli pep

US-10-425-114-65423

Query Match 41.9%; Score 412.5; DB 4; Length 179;

Best Local Similarity 42.3%; Pred. No. 7e-31;

Matches 80; Conservative 28; Mismatches 50; Indels 31; Gaps 3;

Qy 8 OETTSQAPGPGVMCKNLGFFGSGQATMGLCSKCYRETWQAKMTLAEQATA----- 60

Db 12 KEAGCQOPEGPILCINNCGFFGSAATMNCCKHEMIMKQEOAOLASSIDSYNGDN 71

Qy 61 --AQTATTAATAVOPAPVHETKLTCEVERTWIVPHOSSSYQODLVTPAAAPQAVKSSI 118

Db 72 GKGPALATVGAVP-----QVEKTLA-----VQPHVATETSEAAV 109

Qy 119 AAPSPREPNCSSCKRRVGLTGPKRCGNYCALHRYSDKHCTTYDYKAGGEALIAKANP 178

Db 110 IAKAKEGPNRCATCRKRVGLTGFCNRCGNTYCSMHRYSDKHDCQPDYRTAARDALIAKANP 169

Qy 179 LVVAERVK 187

Db 170 VVKAERLKD 178

RESULT 13

US-10-767-701-45710

Sequence 45710, Application US/10767701

Publication No. US20040172684A1

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

FILE REFERENCE: 38-21(53535)B

CURRENT APPLICATION NUMBER: US/10/767,701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 45710

LENGTH: 171

TYPE: PRT

ORGANISM: Sorghum bicolor

FEATURE:

OTHER INFORMATION: Clone ID: SORBI-28MAV03-C6463_1.pep

US-10-767-701-45710

Query Match 41.8%; Score 411.5; DB 4; Length 171;

Best Local Similarity 43.2%; Pred. No. 8.3e-31;

Matches 83; Conservative 27; Mismatches 45; Indels 4;

Qy 8 OETTSQAPGPGVMCKNLGFFGSGQATMGLCSKCYRETWQAKMTLAEQATA----- 66

Db 4 KEAGCQOPEGPILCINNCGFFGSAATMNCCKHEMIMKQEOAOLASSIDSYNGDN 71

Qy 67 -----TAAVOPAPVHETKLTCEVERTWIVPHOSSSYQODLVTPAAAPQAVKSSI 115

Db 58 VNGSGGKGPVIAATVDAVP-----QVEKTLV-----V 98

Qy 116 SSIAPSPREPNCSSCKRRVGLTGPKRCGNYCALHRYSDKHCTTYDYKAGGEALIAKANP 175

Db 99 AAVIPKAKEGPNRCATCRKRVGLTGFCNRCGNTYCSMHRYSDKHDCQPDYRTAARDALIAKANP 158

Qy 176 ANPLVVAERVK 187

Db 159 ANPVVAERLKD 170

RESULT 14

US-10-437-963-123595

Sequence 123595, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 123595

LENGTH: 171

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_26414C.1.pep

US-10-437-963-123595

Query Match 41.4%; Score 407.5; DB 4; Length 171;

Best Local Similarity 45.0%; Pred. No. 2e-30;

Matches 81; Conservative 25; Mismatches 61; Indels 2;

Qy 8 OETTSQAPGPGVMCKNLGFFGSGQATMGLCSKCYRETWQAKMTLAEQATA----- 67

Db 4 KEAGCQOPEGPILCINNCGFFGSAATMNCCKHEMIMKQEOAOLASSIDSYNGDN 71

Qy 68 AAAVOPAPVHETKLTCEVERTWIVPHOSSSYQODLVTPAAAPQAVKSSI 127

Db 62 DSGKEPIIAGHAVAVAQVEKTLV-----VQPAETAGPSEGI 110

Qy 128 RCGSGRKRKRVGLTGPKRCGNYCALHRYSDKHCTTYDYKAGGEALIAKANP 187

Db 111 RCTCRKRVGLTGFCNRCGNTYCSMHRYSDKHDCQPDYRTAARDALIAKANP 170

RESULT 15

US-10-425-115-308732

Sequence 308732, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 16:29:16 ; Search time 25 Seconds
(without alignments)
41.995 Million cell updates/sec

Title: US-10-716-089-20
Perfect score: 984
Sequence: 1 MATERVSEGTSSQAEQPM.....GOEAIKANPLVVAEKVKF 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 32527 seqs, 5584426 residues
Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA New:*
- 1: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB pep.*
 - 2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB pep.*
 - 3: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB pep.*
 - 4: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB pep.*
 - 5: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB pep.*
 - 6: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB pep.*
 - 7: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB pep.*
 - 8: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	8.9	595	US-11-182-946-9	Sequence 9, Appl1
2	80	8.1	2432	US-10-821-234-899	Sequence 899, App
3	79	8.0	259	US-10-131-826A-300	Sequence 300, App
4	79	8.0	259	US-11-182-946-2	Sequence 2, Appl1
5	76.5	7.8	1163	US-11-044-899-2	Sequence 2, Appl1
6	76.5	7.8	1163	US-11-044-899-30	Sequence 30, Appl1
7	72.5	7.4	190	US-11-055-822-538	Sequence 538, App
8	72	7.3	2630	US-11-186-731-2	Sequence 2, Appl1
9	72	7.3	7968	US-11-186-731-5	Sequence 5, Appl1
10	71.5	7.3	502	US-11-021-441-7	Sequence 7, Appl1
11	71.5	7.3	563	US-11-021-441-9	Sequence 9, Appl1
12	71.5	7.3	574	US-11-021-441-15	Sequence 15, Appl1
13	71.5	7.3	581	US-11-021-441-11	Sequence 11, Appl1
14	71.5	7.3	581	US-11-021-441-13	Sequence 13, Appl1
15	71.5	7.3	1035	US-11-021-441-4	Sequence 4, Appl1
16	71	7.2	715	US-10-131-826A-116	Sequence 116, App
17	70.5	7.2	1207	US-10-821-234-1109	Sequence 1109, App
18	70.5	7.2	4419	US-10-821-234-1155	Sequence 1155, App
19	69.5	7.1	1066	US-11-055-822-370	Sequence 370, App
20	69.5	7.1	1066	US-11-055-822-1002	Sequence 1002, App
21	69.5	7.1	1113	US-11-055-822-368	Sequence 368, App
22	69.5	7.1	1113	US-11-055-822-1000	Sequence 1000, App
23	69	7.0	188	US-11-064-774A-117	Sequence 117, App
24	69	7.0	1377	US-10-821-234-1070	Sequence 1070, App
25	68.5	7.0	1005	US-11-113-424-63	Sequence 63, Appl1

25	68	6.9	151	6	US-10-821-234-1378	Seq	App
27	68	6.9	151	6	US-10-528-031-9	Seq	App1
28	68	6.9	409	6	US-10-878-556A-55	Seq	App1
29	68	6.9	943	6	US-10-467-657-5508	Seq	App
30	67.5	6.9	481	6	US-10-467-657-4396	Seq	App
31	67	6.8	985	7	US-11-113-424-61	Seq	App1
32	66.5	6.8	463	6	US-10-131-826A-160	Seq	App
33	66.5	6.8	3507	7	US-11-075-185-7	Seq	App1
34	66	6.7	251	7	US-11-054-515-1538	Seq	App
35	66	6.7	442	6	US-10-821-234-1594	Seq	App
36	66	6.7	712	6	US-10-770-726-66	Seq	App1
37	66	6.7	1178	7	US-11-044-899-29	Seq	App1
38	65	6.6	346	6	US-10-131-826A-432	Seq	App
39	65	6.6	346	6	US-10-967-457-77	Seq	App1
40	65	6.6	358	6	US-10-821-234-878	Seq	App
41	64.5	6.6	1199	6	US-10-821-234-1126	Seq	App
42	64	6.5	465	6	US-10-821-234-1550	Seq	App
43	63.5	6.5	550	6	US-10-878-556A-166	Seq	App
44	63.5	6.5	669	6	US-10-878-556A-87	Seq	App1
45	63	6.4	219	6	US-10-689-742-106	Seq	App

ALIGNMENTS

RESULT 1
US-11-182-946-9
; Sequence 9, Application US/1182946
; Publication No. US20050255100A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Pei
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/11/182,946
; CURRENT FILING DATE: 2005-07-18
; PRIOR APPLICATION NUMBER: US/10/186,643
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US/09/573,986
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-9

Query Match 8.9%; Score 88; DB 7; Length 595.
Best Local Similarity 18.5%; Pred. No. 0.13;
Matches 48; Conservative 30; Mismatches 98; Indels 8;

QY	11	TSQAEQPMCKNLC-----	YQDLYTPAAAPDAQVKSIADE	125
DB	55	TQQCPQRPDKQCEPDYLDADRCCTACTVCSRDILVEKTPCAMNSI		114
QY	30	SOATGGLCKSCYRETV-----MOAKMTALAEQATQAAQTSATPAAVQPI		83
DB	115	STSAVNSCARCFHSVCAGMIVKPGTAKQNTVEPSPVSPACAF		174
QY	84	CEVERTMTVPHOSSS-----YQDLYTPAAAPDAQVKSIADE		125
DB	175	FOAKTPTVSPATSSASTVPVRGTRLMQEAASKLTRAPDS-PSVGRPS		233
QY	126	PNRQSCCKRGVGLGFKRCGN-----LYCALHRSYDKHTCTYD-----		167
DB	234	PEGSDCKQCEPDYLDADRCCTACTVCSRDILVEKTPCAMNSRICE		293
QY	168	AGOEAIAKANPLVVAEKVK 187		
DB	294	TNSCARCVPIICAAETVTK 313		

```

RESULT 2
US-10-821-234--899
Sequence 899, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmani, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821.234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: pc_seq_genes version 1.0
SEQ ID NO 899
LENGTH: 2432
TYPE: PRN
ORGANISM: Homo sapiens
US-10-821-234--899

Query Match      8.1%; Score 80; DB 6; Length 2432;
Best Local Similarity 21.8%; Pred. No. 3.7;
Matches 37; Conservative 37; Mismatches 78; Indels 18; Gaps 6;

QY      5 RVSGETTSQAPEGPMCKNLGPGSQATWGLGSKCYRETYM---QAKMTALAQQTAA 61
Db      984 RIARPRRIADR-PMLASRSMMSTYAERSMMSSYSRSMMSYERSMSPMAERSMMS 10422
QY      62 QATSATAAAVOP-PAPVHETKLTCEVERTMTVPHGSSSYQODLVTPAA-----AAPQA 113
Db      1043 YERSMMSAYERSMMSPMARSMMMSAYERSM-----SAYERSMMSPMADRSMMMSGADRS 10974
QY      114 VKGSIAAPSREPNNRCSCRRKRGVLGTGKKCGCNLYCALHHYSDDHTCTY 163
Db      1098 MMSSYSIADDRSMMSY-SAADRSMMSYSTADRSMMMSADSYTSDTYD 1146

RESULT 3
US-10-131-826A-300
Sequence 300, Application US/10131826A
Publication No. US20050245730A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RIC128
CURRENT APPLICATION NUMBER: US/10/131.826A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26

```

```

PRIORITY APPLICATION NUMBER: 60/059113
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059115
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059117
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059122
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059184
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059263
PRIORITY FILING DATE: 1997-09-18
PRIORITY APPLICATION NUMBER: 60/059352
PRIORITY FILING DATE: 1997-09-19
PRIORITY APPLICATION NUMBER: 60/059588
PRIORITY FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 300
LENGTH: 259
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-826A-300

Query Match
Best Local Similarity 8.0%; Score 79; DB 6; Length 259;
Matches 34; Conservative 19; Mismatches 62; Indels 6;

6 VSQRTTSQAPRG-----PWCNCLGCFRFSQATMGLCS-----KCYR
103 MTRITVCCCKEGTFRNENSPFMCCKRCRCBSGEGVQVSNCTSMWDIOQVE
Matches 34; Conservative 19; Mismatches 62; Indels 6;

6 LAEQATQAQATSATA-----AVQPPAPVHETKLTCE-----VER
163 AAEETMNTSPGTAPAAAEETMNTSPGTAPAAAEETMNTSPGTAPAAAE
100 QQDLVTPAAAAAPQVKSSIAAPSRPEENRCGSC 132
214 ---TTSPGTAPAAAEETMNTSPGTAPASSHYISC 243

Db

RESULT 4
US-11-182-946-2
Sequence 2, Application US/11182946
Publication No. US20050255100A1
GENERAL INFORMATION:
APPLICANT: Wei, Ying-Fei
APPLICANT: Ni, Jian
APPLICANT: Gentz, Retner
ATTORNEY: Ruben, Steven
TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
FILE REFERENCE: 1488.1280004
CURRENT APPLICATION NUMBER: US/11/182,946
CURRENT FILING DATE: 2005-07-18
PRIOR APPLICATION NUMBER: US/10/186,643
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US/09/573,986
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 259
TYPE: PRT
ORGANISM: Homo sapiens
US-11-182-946-2

Query Match
Best Local Similarity 8.0%; Score 79; DB 7; Length 259;
Matches 34; Conservative 19; Mismatches 62; Indels 6;

6 VSQRTTSQAPRG-----PWCNCLGCFRFSQATMGLCS-----KCYR
103 MTRITVCCCKEGTFRNENSPFMCCKRCRCBSGEGVQVSNCTSMWDIOQVE
Matches 34; Conservative 19; Mismatches 62; Indels 6;

6 VSQRTTSQAPRG-----PWCNCLGCFRFSQATMGLCS-----KCYR
103 MTRITVCCCKEGTFRNENSPFMCCKRCRCBSGEGVQVSNCTSMWDIOQVE
Matches 34; Conservative 19; Mismatches 62; Indels 6;

```



```

Qy      28  FGSQATMGL-----CSKCRRE--TWQAKTALAEQATQAAQTSATPAAVQPPAPVHE 79
Db      8  FGLAATAALSAALVACSPDPHOQDSPVQRTNEILTTSCQPTSSASSITSSATTSSAPVEE 67

Qy      80  TKLTCEVERTMIVPHQSSSYQODLVTPPAAAPQAVKSIAPSPRPNNRCG 130
Db      68  -DVEIVVSPALVDSGEVTFEIRISGLDPEGCVYAAICQSVANRGNVPVPSCTG 117

RESULT 8
US-11-186-731-2
; Sequence 2, Application US/11186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MP12001-047P1RCP1 (M)
; CURRENT APPLICATION NUMBER: US/11/186,731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-731-2

Query Match      7.3%; Score 72; DB 7; Length 2630;
Best Local Similarity 26.5%; Pred. No. 24;
Matches 30; Conservatve 15; Mismatches 36; Indels 32; Gaps 5;

Qy      24  LCGFPGSQATMGLCSKCYRETWQAKTALAEQATQAAQTSATPAAVQPPAPVHETKLT 83
Db      1564  LKGGTALGALGGL-----REPMEHRV--LEBEARREQAT-----LLAKAPSFETILR 1610

Qy      84  CEVERTMIVPHQSSSYQODLVTPPAAAPQAVKSIAPSPRPNN--RCGSCKR 134
Db      1611  LPAAGTHLAPHSHSLND-----SPTSRPSEACGGBAQR 1646

RESULT 9
US-11-186-731-5
; Sequence 5, Application US/11186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MP12001-047P1RCP1 (M)
; CURRENT APPLICATION NUMBER: US/11/186,731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 7968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-731-5

Query Match      7.3%; Score 72; DB 7; Length 7968;
Best Local Similarity 26.5%; Pred. No. 81;

```

Matches	30; Conservative	15; Mismatches	36; Indels	
QY	24	LCGFGSOATWGLGCSKCYRETVMQAKMTALAEQATQAAQATSATAAAC		83
Db	6902	LKGGYIALPGL-----REPLMEHRV--LEEBAAREQAT-----LL		6948
QY	84	CEVERTMIVPHOSSYQODLVTPAAAPQAVKSSIAAPSRPEPN--RCG		
Db	6949	LPASGTHLAPGSHSLERD-----SPSTPRPSSEACG		
RESULT 10				
US-11-021-441-7				
/ Sequence 7, Application US/11021441				
/ Publication No. US20050249748A1				
/ GENERAL INFORMATION:				
/ APPLICANT: DUBENSKY, Thomas W., Jr.				
/ APPLICANT: PORTNOY, Daniel A.				
/ APPLICANT: LUCKETT, William S., Jr.				
/ APPLICANT: COOK, David N.				
/ TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,				
/ TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND				
/ TITLE OF INVENTION: THEREOF				
/ FILE REFERENCE: 282172003900				
/ CURRENT FILING DATE: 2004-12-23				
/ PRIOR APPLICATION NUMBER: US 60/616,750				
/ PRIOR FILING DATE: 2004-10-06				
/ PRIOR APPLICATION NUMBER: US 60/615,287				
/ PRIOR FILING DATE: 2004-10-01				
/ PRIOR APPLICATION NUMBER: US 60/599,377				
/ PRIOR FILING DATE: 2004-08-05				
/ PRIOR APPLICATION NUMBER: PCT/US2004/33881				
/ PRIOR FILING DATE: 2004-07-23				
/ PRIOR APPLICATION NUMBER: US 10/883,599				
/ PRIOR FILING DATE: 2004-06-30				
/ PRIOR APPLICATION NUMBER: US 60/556,744				
/ PRIOR FILING DATE: 2004-03-26				
/ NUMBER OF SEQ ID NOS: 129				
/ SOFTWARE: FastSeq for Windows Version 4.0				
/ SEQ ID NO 7				
/ LENGTH: 502				
/ TYPE: PRT				
/ ORGANISM: Homo sapiens				
US-11-021-441-7				
Query Match	7.3%; Score 71.5; DB 7; Length 50			
Best Local Similarity	22.8%; Pred. No. 4.3; Indels			
Matches	34; Conservative	18; Mismatches	62; Indels	
QY	15	PEGPYMC-----KNLC-----GFGSOATWGLGCSKCYRETVM-QQAK		60
Db	232	PIGGLCGAGYEKVEDACQACSPGFKFEASBSPLCEPHTLPSPEGA'		291
QY	61	AAQTSATAAVQPPAPVY---ETLTCVEKTMIVPHOSSYQODLVTP:		117
Db	292	APQDPASMPCTRPSPAPHYLTAVAGAGAYELRW--TPPDGSGREDI---		340
QY	118	IAAPSRPEPNRCGSCRKRV-----GLT 139		
Db	341	TCEQCPESGEGCPCEASVRYSEPHGLT 369		
RESULT 11				
US-11-021-441-9				
/ Sequence 9, Application US/11021441				
/ Publication No. US20050249748A1				
/ GENERAL INFORMATION:				
/ APPLICANT: DUBENSKY, Thomas W., Jr.				
/ APPLICANT: PORTNOY, Daniel A.				
/ APPLICANT: LUCKETT, William S., Jr.				
/ APPLICANT: COOK, David N.				
/ TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,				

```

; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-11-021-441-9

```

```

Query Match          7.3%; Score 71.5; DB 7; Length 563;
Best Local Similarity 22.8%; Pred. No. 4.8;
Matches 34; Conservative 18; Mismatches 62; Indels 35; Gaps 7;

```

```

QY 15 PEGPVMC-----KNLC-----GFGSQATMGLCSKCYRETV--MQAKMTALAEQATQA 60
DB 293 PIGQCLCGAGYKVEDACQACSPGFFKFEASESPCLCEPHTLPSPPEG 352
QY 61 AQAATATAAAGVAPPAVH---ETKLTCEVERTMIIPHQSSSYOODLVTPAALAPQAVKSS 117
DB 353 APQDPAISMCTRPSPAPHYLTAVGMGAKVELRW--TPPDGSGREDI-----VTSV 401
QY 118 IAAPSRPEPNRCGSCRKRV-----GLT 139
DB 402 TCEQCWPESGECGPCEASVRYSEBPHGLT 430

```

```

RESULT 12
US-11-021-441-15
; Sequence 15, Application US/11021441
; Publication No. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 15
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-11-021-441-15

```

```

Query Match          7.3%; Score 71.5; DB 7; Length 574
Best Local Similarity 22.8%; Pred. No. 4.9;
Matches 34; Conservative 18; Mismatches 62; Indels 7;

```

```

QY 15 PEGPVMC-----KNLC-----GFGSQATMGLCSKCYRETV--MQAK 60
DB 294 PIGQCLCGAGYKVEDACQACSPGFFKFEASESPCLCEPHTLPSPPEG 353
QY 61 AQAATATAAAGVAPPAVH---ETKLTCEVERTMIIPHQSSSYOODLVTP 117
DB 354 APQDPAISMCTRPSPAPHYLTAVGMGAKVELRW--TPPDGSGREDI-- 402
QY 118 IAAPSRPEPNRCGSCRKRV-----GLT 139
DB 403 TCEQCWPESGECGPCEASVRYSEBPHGLT 431

```

```

RESULT 13
US-11-021-441-11
; Sequence 11, Application US/11021441
; Publication No. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-11-021-441-11

```

```

Query Match          7.3%; Score 71.5; DB 7; Length 581
Best Local Similarity 22.8%; Pred. No. 5;
Matches 34; Conservative 18; Mismatches 62; Indels 7;

```

```

QY 15 PEGPVMC-----KNLC-----GFGSQATMGLCSKCYRETV--MQAK 60
DB 301 PIGQCLCGAGYKVEDACQACSPGFFKFEASESPCLCEPHTLPSPPEG 360
QY 61 AQAATATAAAGVAPPAVH---ETKLTCEVERTMIIPHQSSSYOODLVTP 117
DB 361 APQDPAISMCTRPSPAPHYLTAVGMGAKVELRW--TPPDGSGREDI-- 409

```

QY 118 IAAPSRPEPNRGSCRKRV-----GLT 139
DB 410 TCBCQWPSGCGPCBASVRYSEPPHGLT 438

RESULT 14

US-11-021-441-13
; Sequence 13, Application US/11021441
; Publication NO. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
; US-11-021-441-13

Query Match 7.3%; Score 71.5; DB 7; Length 581;

Best Local Similarity 22.8%; Pred. No. 5;
Matches 34; Conservative 18; Mismatches 62; Indels 35; Gaps 7;

QY 15 PEGPVMC-----KNLC-----GFGSQATWGLCSKCYRETV--MOAKTALAEQATQA 60
DB 301 PIGQCLCOAGYEKVEDACQACSPGFKFEASBSPCLCEPHTLPSPEGATSCCEGFFR 360
QY 61 AQTATATAAVQPPAPVH---ETKLTCVEKRTMIVPHOSSYQODLVTPAAAAQAVKSS 117
DB 361 APQDPAHMPCTRPSPAPHYLTAVGMGAKVELRW--TPQDSGGREDI-----YYSV 409
QY 118 IAAPSRPEPNRGSCRKRV-----GLT 139
DB 410 TCBCQWPSGCGPCBASVRYSEPPHGLT 438

RESULT 15

US-11-021-441-4
; Sequence 4, Application US/11021441
; Publication NO. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 282172003900

; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
; US-11-021-441-4

Query Match 7.3%; Score 71.5; DB 7; Length 10

Best Local Similarity 22.8%; Pred. No. 9.5;
Matches 34; Conservative 18; Mismatches 62; Indels 7;

QY 15 PEGPVMC-----KNLC-----GFGSQATWGLCSKCYRETV--MOAK 60
DB 315 PIGQCLCOAGYEKVEDACQACSPGFKFEASBSPCLCEPHTLPSPEGA 374
QY 61 AQTATATAAVQPPAPVH---ETKLTCVEKRTMIVPHOSSYQODLVTF 117
DB 375 APQDPAHMPCTRPSPAPHYLTAVGMGAKVELRW--TPQDSGGREDI--- 423
QY 118 IAAPSRPEPNRGSCRKRV-----GLT 139
DB 424 TCBCQWPSGCGPCBASVRYSEPPHGLT 452

Search completed: December 8, 2005, 16:36:29
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 8, 2005, 16:33:01 ; Search time 3238 Seconds
(without alignment)
3300.361 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATERVSGEITSGAPEGPMV.....GGEAIKANPLVAKKVKVF 188

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10716089/runat.05122005.094343.13141/app.query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=sp2n.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human0.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716089@cgn.1.1.7415@runat.05122005.094343.13141 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_in:.*
3: gb_env:.*
4: gb_com:.*
5: gb_ov:.*
6: gb_pal:.*
7: gb_ph:.*
8: gb_pr:.*
9: gb_ro:.*
10: gb_srs:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_vl:.*
14: gb_hcg:.*
15: gb_pl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984	100.0	804	6	AR452950 Sequence
2	984	100.0	804	6	AX288141 Sequence
3	957	97.3	1531	6	AR452942 Sequence

4	957	97.3	1531	6	AX288133	3e
5	419.5	42.6	999	15	AK104562	sat
6	419.5	42.6	1000	15	AK104175	AK3
7	419.5	42.6	1009	15	AK104605	sat
8	419.5	42.6	1021	15	AK060008	AK0
9	419.5	42.6	1209	15	AK069720	sat
10	419.5	42.6	110000	15	AP008212_240	AK0
11	419.5	42.6	110000	15	AP008212_241	Con
12	419.5	42.6	158321	15	AP003626	Con
13	419.5	42.6	163568	15	AP003711	AP0
14	416.5	42.3	522	6	AK654090	AK5
15	416.5	42.3	991	15	AK061911	AK0
16	416.5	42.3	1000	15	AK067403	AK0
17	416.5	42.3	110000	15	AP008208_033	Con
18	416.5	42.3	142732	15	AP004184	AP0
19	415.5	42.2	1132	15	AY345599	AY3
20	415	42.2	1020	15	BT016379	BT0
21	413	42.0	1162	15	AY515607	AY5
22	409.5	41.6	1026	15	BT014337	BT0
23	407.5	41.4	117676	15	AP006427	AP0
24	400.5	40.7	112882	14	AC161406	AC1
25	390.5	39.7	100208	15	AC155881	AC1
26	387.5	39.4	1032	15	BT001984	BT0
27	387	39.3	941	15	BT017601	BT0
28	386	39.2	798	6	AK653271	AK5
29	386	39.2	854	15	AY377427	AY3
30	386	39.2	110000	15	AP008209_327	Con
31	386	39.2	162497	15	AC090871	AC0
32	385.5	39.2	125195	15	AC154034	AC1
33	385	39.1	1158	15	AK102503	AK1
34	384.5	39.1	486	6	AK654221	AK5
35	384.5	39.1	727	15	AY282740	AY2
36	384.5	39.1	887	15	AK121813	AK1
37	384.5	39.1	110000	15	AP008213_036	Con
38	384.5	39.1	110000	15	AP008213_037	Con
39	384.5	39.1	151986	15	AP005840	AP0
40	383.5	39.0	102653	14	AP007400	AP0
41	382.5	38.9	898	15	AY056445	AY0
42	379	38.5	522	15	BT015115	BT0
43	379	38.5	105807	15	AC006085	AC0
44	371	37.7	2786	5	CR761102	CR7
45	369.5	37.6	2161	5	BC081266	BC0

ALIGNMENTS

RESULT 1
AR452950
LOCUS AR452950 804 bp DNA linear
DEFINITION Sequence 12 from patent US 6677504.
ACCESSION AR452950
VERSION AR452950.1 GI:42685097
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 804)
AUTHORS da Costa e Silva,O., Bohnert,H.J., van Thiel,N., a
TITLE Transcription factor stress-related proteins and me
JOURNAL Patent: US 6677504-A 12 13-JAN-2004;
BASF Plant Science GmbH; Ludwigshafen;
DEX;

FEATURES
source Location/Qualifiers
1..804
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 6,32e-61 Length: 804
Score: 984.00 Matches: 188
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-716-089-20 (1-188) x AR452950 (1-804)			
QY	1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet	20	
DB	134 ATGGCCACCGAGCGTGTCTCAGAGACGACCTGGCAGGCGCTCGACATTATG	193	
QY	21 CysLysAsnLeuCyseGlyPhePheGlySerGlnAlaThrMetGlyLeuCyseSerLysCyse	40	
DB	194 TGCAGAACCTTTGCGGCTTTCTTCGGCAGCCAACTGACATGGGGTTGGCTCGAAGTGC	253	
QY	41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla	60	
DB	254 TACCGAGAGACAGTCATGCAAGCAAGATGACGGCTTTAGCTGACCAAGCCCTCAGGCT	313	
QY	61 AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaProValHisGluThr	80	
DB	314 GCTCAGGCGACATCTGCCACAGCTGCTGTTCAAGCCCCCGCTCCTGATCAGAGACC	373	
QY	81 LysLeuThrCyseGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln	100	
DB	374 AAGCTCACATGCGAGGTTGAGAGAACATGATTGGCCGATCAATCTTCAGCTATCAA	433	
QY	101 GlnAspLeuValThrProAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla	120	
DB	434 CAAAGCCTGGTTACCCCGCTCAGCTGCCCTCAGGCGAGTAAAGTCTTATCTCAGCT	493	
QY	121 ProSerArgProGluProAsnArgCyseGlySerCyseArgLysArgValGlyLeuThrGly	140	
DB	494 CCTCTAGACCCGAGCCCAATCGATCGGATCTTCGAGGAACCGTTCGATGACAGGA	553	
QY	141 PheLysCyseArgCyseGlyAsnLeuTyrCyseAlaLeuHisArgTyrSerAspLysHisThr	160	
DB	554 TTTAAGTGTGCTGTGGCAACTCTACTGCTTAACTCGGTACTCGGACAAACACT	613	
QY	161 CysThrTyrAspTyrLysAlaAlaGluGlnGluAlaIleAlaLysAlaAsnProLeuVal	180	
DB	614 TGCACATATGACTACAAAGCCGAGGCGAGGAGCATTTGCAAAAGCTTATCTCTTGC	673	
QY	181 ValAlaGluLysValValLysPhe	188	
DB	674 GTGGCCGAGAGGTTGTCAAGTTT	697	
RESULT 2			
LOCUS	AX288141	804 bp	DNA linear PAT 21-NOV-2001
DEFINITION	Sequence 12 from Patent WO0177311.		
ACCESSION	AX288141		
VERSION	AX288141.1	GI:17049843	
KEYWORDS	Phycomitrella patens		
SOURCE	Phycomitrella patens		
ORGANISM	Phycomitrella patens		
REFERENCE	1 da Costa Silva,O., Bohner,H.U., van Thiel,N. and Chen,R. Transcription factor stress-related proteins and methods of use in plants Patent: WO 0177311-A 12 18-OCT-2001;		
JOURNAL	BASF Plant Science GmbH (DE)		
FEATURES	source 1..804 /organism="Phycomitrella patens" /mol_type="unassigned DNA" /db_xref="taxon:3218"		
ORIGIN			
Alignment Scores:	6.32e-61	Length:	804
Pred. No.:	984.00	Matches:	188
Score:			

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-716-089-20 (1-188) x AX288141 (1-804)			
QY	1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluG		20
DB	134 ATGGCCACCGAGCGTGTGTCTCAGAGACGACCTGGCAGGCGCTCGAGG		193
QY	21 CysLysAsnLeuCyseGlyPhePheGlySerGlnAlaThrMetGlyLeuCy		40
DB	194 TGCAGAACCTTTGCGGCTTTCTTCGGCAGCCAACTGACATGGGGTTGT		253
QY	41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGln		60
DB	254 TACCGAGAGACAGTCATGCAAGCAAGATGACGGCTTTAGCTGAGCAG		313
QY	61 AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaProVal		80
DB	314 GCTCAGGCGACATCTGCCACAGCTGCTGTTCAAGCCCCCGCTCCTCG		373
QY	81 LysLeuThrCyseGluValGluArgThrMetIleValProHisGlnSerS		100
DB	374 AAGCTCACATGCGAGGTTGAGAGAACATGATTGGCCGATCAATCTT		433
QY	101 GlnAspLeuValThrProAlaAlaAlaProGlnAlaValLysSerS		120
DB	434 CAAAGCCTGGTTACCCCGCTCAGCTGCCCTCAGGCGAGTAAAGTCTC		493
QY	121 ProSerArgProGluProAsnArgCyseGlySerCyseArgLysArgValG		140
DB	494 CCTCTAGACCCGAGCCCAATCGATCGGATCTTCGAGGAACCGTTCG		553
QY	141 PheLysCyseArgCyseGlyAsnLeuTyrCyseAlaLeuHisArgTyrSerA		160
DB	554 TTTAAGTGTGCTGTGGCAACTCTACTGCTTAACTCGGTACTCGG		613
QY	161 CysThrTyrAspTyrLysAlaAlaGluGlnGluAlaIleAlaLysAlaA		180
DB	614 TGCACATATGACTACAAAGCCGAGGCGAGGAGCATTTGCAAAAGCTT		673
QY	181 ValAlaGluLysValValLysPhe	188	
DB	674 GTGGCCGAGAGGTTGTCAAGTTT	697	
RESULT 3			
LOCUS	AR452942	1531 bp	DNA linear 2004
DEFINITION	Sequence 4 from patent US 6677504.		
ACCESSION	AR452942		
VERSION	AR452942.1	GI:42685089	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1531) da Costa e Silva,O., Bohner,H.U., van Thielen,N. a Transcription factor stress-related proteins and me plants Patent: US 6677504-A 13-JAN-2004; BASF Plant Science GmbH; Ludwigshafen; DEX;		
JOURNAL			
FEATURES	Location/Qualifiers 1..1531 /organism="unknown" /mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:	1.03e-58	Length:	1531
Pred. No.:	957.00	Matches:	187
Score:			

Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.26% Indels: 1
DB: 6 Gaps: 0

US-10-716-089-20 (1-188) x AR452942 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
DB 655 ATGGCCACCGAGCGTGTCTCAGAGAGACGACCTCCAGGCCCTTAGGGTCCAGTTAAG 714
QY 21 CysLysAsnLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerLysCys 40
DB 715 TGCMAAACCTTTGGCGCTTCTTCGCGACGCCAAGCTAACATGGGTTGCTGAAAGTGC 774
QY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGAGACAGTCATGACGCG-AAAGATGACGGCTTTAGCTGACAGCACTCAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHISGluThr 80
DB 834 GCTCAGCGCAGCATCTGCCACAGCTGCTGCTTCACCCCCCTCTGTACATGAGACC 893
QY 81 LysLeuThrCyGlyGluValGluArgThrMetIleValProHISGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGCGAGGTGAGAGAACATGATTTGCGCGATCAATCTTCCAGCTATCA 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla 120
DB 954 CAAGACCTGGTTACCCCGCTGACGCTGCCCTCAGCGAGTGAAGCTCTATCCGAGCT 1013
QY 121 ProSerArgProGluProAlaAsnArgCyGlyLysSerCyArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGCCCAATCATGCGATCTTGAGAGAGCGTTGGATTTGACAGCA 1073
QY 141 PheLysCyArgCyGlyAsnLeuTyrCyAlaLeuHisArgTyrSerAspLysHisThr 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTACTGCGCTTTACATCGGTACTCGACAAACACT 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal 180
DB 1134 TGCACATATGACTACAAAGCCGACGAGCAAGACGATTGCCAAACTATCTCTTGTTC 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 4
AX288133 1531 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 4 from Patent WO0177311.
ACCESSION AX288133
VERSION AX288133.1 GI:17049835
KEYWORDS
SOURCE . Physcomitrella patens
ORGANISM Physcomitrella patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE 1 da Costa Silva, O., Bohmert, H.J., van Thiel, N. and Chen, R.
AUTHORS Transcription factor stress-related proteins and methods of use in
TITLE plants
JOURNAL Patent: WO 0177311-A 4 18-OCT-2001;
FEATURES
source BASF Plant Science GmbH (DE)
Location/Qualifiers
1..1531
/organism="Physcomitrella patens"
/mol_type="unassigned DNA"
/db_xref="taxon:3218"

Alignment Scores:
Pred. No.: 1.03e-58 Length: 1531

Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.26% Indels: 1
DB: 6 Gaps: 0

US-10-716-089-20 (1-188) x AX288133 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluG 20
DB 655 ATGGCCACCGAGCGTGTCTCAGAGAGACGACCTCCAGGCCCTTAGAG 714
QY 21 CysLysAsnLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeu 40
DB 715 TGCMAAACCTTTGGCGCTTCTTCGCGACGCCAAGCTAACATGGGTTG 774
QY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGln 60
DB 775 TACCGAGAGACAGTCATGACGCG-AAAGATGACGGCTTTAGCTGACCA 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaPro 80
DB 834 GCTCAGCGCAGCATCTGCCACAGCTGCTGCTTCAGCCCCCTCTCTG 893
QY 81 LysLeuThrCyGlyGluValGluArgThrMetIleValProHISGlnSer 100
DB 894 AAGCTCACATGCGAGGTGAGAGAACATGATTTGCGCGATCAATCTT 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSer 120
DB 954 CAAGACCTGGTTACCCCGCTGACGCTGCCCTCAGCGAGTGAAGCTCT 1013
QY 121 ProSerArgProGluProAlaAsnArgCyGlyLysSerCyArgLysArgValG 140
DB 1014 CCTCTAGACCCGAGCCCAATCATGCGATCTTGAGAGAGCGTTT 1073
QY 141 PheLysCyArgCyGlyAsnLeuTyrCyAlaLeuHisArgTyrSer 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTACTGCGCTTTACATCGGTACTCG 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAla 180
DB 1134 TGCACATATGACTACAAAGCCGACGAGCAAGACGATTGCCAAAGCT 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 5
AK104562 999 bp mRNA linear
LOCUS AK104562
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:O
insert sequence.
ACCESSION AK104562
VERSION AK104562.1 GI:32989771
KEYWORDS FLI_CDNA; oligo capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; I
Euphorbiaceae; Oryzaeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Inst
Agrobiological Sciences Rice Full-Length cDNA Proj
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Do
Kishimoto, N., Yarak, J., Ishikawa, M., Yamada, H., Oo
Kojima, K., Nemiki, T., Ohneda, E., Yahagi, W., Suzuki,
Onseuki, K., Shisuniki, T., Foundation of Advancement
Science Genome Sequencing & Analysis Group.; Oono,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xi
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., N
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., M

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
Japonica rice
Science 301 (5631), 376-379 (2003)
JOURNAL
PUBMED
REFERENCE
AUTHORS
2 (bases 1 to 1009)
12869764
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoaka, T.,
Hori, P., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Maeda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Naniki, T., Narikawa, R., Niihara, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and
Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from Japonica
rice.
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Naniki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Maeda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K., and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoaka, T., Hori, P., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A., and Hayashizaki, Y.
Location/Qualifiers
1. 1009
/organism="Oryza sativa (Japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="006-309-A12"

ORIGIN
Alignment Scores:
Pred. No.: 6.96e-21 Length: 1009
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2
US-10-716-089-20 (1-188) x AK104605 (1-1009)
OY 8 GcngluthrThrsrGlnAlaProgluGlyProValMetCysLysAsnL
:::|||||
DB 159 AAGAGAGCTGATGATCCGAGCGAGCGGAGGCCGATCTTCCTTCATCAATTA
OY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerSerCysTyrAlaGluT
:::|||||
DB 219 TTGGCAGTGGCGCTACCATGATGATGCTGCTCAAGTGCACAGAGAGA
OY 48 AlAlaYsMetThrAlaLeuAlaGluGlnAlaThrGlnAlaGlnAlaT
:::|||||
DB 279 CAGAGCAGCGCCAGAGCTGGCA-----GCCCTCTTCATCCAGCATTC
OY 68 AlAlaAlaValAlaGlnProProAlaProValAlaGluThrLysLeuThrC
:::|||||
DB 333 GATTCGGGAGAGAACCAATTAATTCGTGCTGATCAGCTGAGTAAAGCTGTTG
OY 88 ArgThrMetLeuValProHisGlnSerSerSerTyrGlnGlnAspLeuV
:::|||||
DB 393 GTGAAGACGCTTGT-----G
OY 108 AlAlaAlaProGlnAlaValAlaValSerSerLeuAlaAlaProSerArgP
:::|||||
DB 420 GAGATTGCTGCTCTTATGAGAGGGGCTCAGCTGAACCCAGGAGAGG
OY 128 ArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysA
:::|||||
DB 480 CGGCTCCTCACTTGTTCGGAAGAGGGTGTCTTACCGGATTCACATGCTCC
OY 148 LeuTyrCysAlaLeuHisArgTyrSerAspLysHisThrCysThrTyrA
:::|||||
DB 540 TTGTACTGCGGCAAGCAACCGCTAATTCGATTAAGATGATGCTCCAGTTTG
OY 168 AlAlaGlyGlnGlnAlaLeuAlaLeuAlaAlaAspProLeuValValAlaGluL
:::|||||
DB 600 GCTGCTAGGATGCTATTGCTCAAGGCTTAATCCGGTGTGAAGGCGAGAG
RESULT 8
AK060008 1021 bp mRNA linear 2003
LOCUS Oryza sativa (Japonica cultivar-group) cDNA clone:0:
DEFINITION
Accession
AK060008
VERSION AK060008.1 GI:32970026
KEYWORDS FLI cDNA; oligo-capping.
SOURCE Oryza sativa (Japonica cultivar-group)
ORGANISM Oryza sativa (Japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta
Spermatophytes; Magnoliophyta; Liliopsida; Poales; P
Ehretidoideae; Oryzaceae; Oryza.
REFERENCE
1 The Rice Full-length cDNA Consortium, National Inst
Agrobiological Sciences Rice Full-length cDNA Proje
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Do
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Oo
Kojima, K., Naniki, T., Ohneda, E., Yahagi, W., Suzuki,
Ohtsuki, K., Shishiki, T., Foundation of Advancement
Science Genome Sequencing & Analysis Group, Otsu,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda,
Kurosaki, T., Kodama, T., Maeda, H., Kobayashi, M., Xi
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., N
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., M.

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

COMMENT

TITLE

JOURNAL

FEATURES

source

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Otsu, Y., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice Science 301 (5631), 376-379 (2003)

12869784

2 (bases 1 to 1209)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujiwara, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, J., Iida, J., Ikeda, R., Imamura, K., Imotani, K., Ishidaki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kuroaki, T., Kusumegi, T., Li, C., Lu, M., Maeda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nishikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Otsu, N., Ota, Y., Otsu, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28k full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujiwara, T., Ikeda, R., Ishidaki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kuroaki, T., Kusumegi, T., Lu, M., Maeda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, T., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohta, Y., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1..1209

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultiar="Nippondare"

/db_xref="taxon:39947"

/clone="J023026K20"

ORIGIN

Alignment Scores:

Pred. No.: 8,46e-21 Length: 1209

Score: 419.50 Matches: 82

Percent Similarity: 60.00% Conservative: 26

Best Local Similarity: 45.56% Mismatches: 59

Query Match: 42.63% Indels: 13

DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AK069720 (1-1209)

Qy 8 GlnGlnThrThrSerGlnAlaProGluGlyProValMetCysValAsnL

Db 157 AAGAGATGGATGATCCAGCAGCCGAGAGCCGATCTTCATCAATAT

Qy 28 PheGlySerGlnAlaThrMetGlyLeuCysSerCysValTAAAGLTI

Db 217 TTTCGAGTGGCGCTACCATGATGATGCTCAAGTCCCAAGAGAA

Qy 48 AlAlaYmecthrAlaLeuAlaGluGlnAlaThrGlnAlaGlnAlaT

Db 277 CAGGAGCAGGCCAAGCTGGCA-----GCTCTCTATCCAGCATTC

Qy 68 AlAlaAlaValGlnProProAlaProValHisGluThrIlyLeuThrC

Db 331 GATTCGGGAGAACCAATATTGCTGCTGCTGCTGCTGCTGCTGCTG

Qy 88 ArgThrMetIleValProHisGlnSerSerSerIyGlnGlnApleuV

Db 391 GTGAAGACGCTTGT-----G

Qy 108 AlAlaAlaProGlnAlaValAlaValSerSerIleAlaAlaProSerArgP

Db 418 GAGATTGCTGGCCCTTAGTGAAGGGCTCAGGTGAACCCAGGAGGG

Qy 128 ArgCysGlySerCysArgIlyArgValGlyLeuThrGlyPheIlyCysA

Db 478 CGGAGTCCACTTGTCCGAAAGAGGTTGTTCTTACCGGATTCACCTGCC

Qy 148 LeuIlyCysAlaLeuHisIlyArgIlySerApleyHisIlyThrCysThrIyA

Db 538 TTGTAATCGCGCAAGCAACCGCTATTCCGTAAGCATGACTCCAGTTTG

Qy 168 AlAgIyGlnGlnAlaIleAlaIleValAlaAsnProLeuValAlaGluL

Db 598 GCTGCTAGGGATGCTATTGCTCAAGGCTTAATCCGTTGTTGAAGCGGAGA

RESULT 10

AP008212_240

WPCOMMENT

Sequence split into 308 fragments LOCUS AP008212 Accession A1

Fragment Name	Begin	End
AP008212_000	1	110000
AP008212_001	100001	210000
AP008212_002	200001	310000
AP008212_003	300001	410000
AP008212_004	400001	510000
AP008212_005	500001	610000
AP008212_006	600001	710000
AP008212_007	700001	810000
AP008212_008	800001	910000
AP008212_009	900001	1010000
AP008212_010	1000001	1110000
AP008212_011	1100001	1210000
AP008212_012	1200001	1310000
AP008212_013	1300001	1410000
AP008212_014	1400001	1510000
AP008212_015	1500001	1610000
AP008212_016	1600001	1710000
AP008212_017	1700001	1810000
AP008212_018	1800001	1910000
AP008212_019	1900001	2010000

AP008212_020 2000001 2110000
AP008212_021 2100001 2210000
AP008212_022 2200001 2310000
AP008212_023 2300001 2410000
AP008212_024 2400001 2510000
AP008212_025 2500001 2610000
AP008212_026 2600001 2710000
AP008212_027 2700001 2810000
AP008212_028 2800001 2910000
AP008212_029 2900001 3010000
AP008212_030 3000001 3110000
AP008212_031 3100001 3210000
AP008212_032 3200001 3310000
AP008212_033 3300001 3410000
AP008212_034 3400001 3510000
AP008212_035 3500001 3610000
AP008212_036 3600001 3710000
AP008212_037 3700001 3810000
AP008212_038 3800001 3910000
AP008212_039 3900001 4010000
AP008212_040 4000001 4110000
AP008212_041 4100001 4210000
AP008212_042 4200001 4310000
AP008212_043 4300001 4410000
AP008212_044 4400001 4510000
AP008212_045 4500001 4610000
AP008212_046 4600001 4710000
AP008212_047 4700001 4810000
AP008212_048 4800001 4910000
AP008212_049 4900001 5010000
AP008212_050 5000001 5110000
AP008212_051 5100001 5210000
AP008212_052 5200001 5310000
AP008212_053 5300001 5410000
AP008212_054 5400001 5510000
AP008212_055 5500001 5610000
AP008212_056 5600001 5710000
AP008212_057 5700001 5810000
AP008212_058 5800001 5910000
AP008212_059 5900001 6010000
AP008212_060 6000001 6110000
AP008212_061 6100001 6210000
AP008212_062 6200001 6310000
AP008212_063 6300001 6410000
AP008212_064 6400001 6510000
AP008212_065 6500001 6610000
AP008212_066 6600001 6710000
AP008212_067 6700001 6810000
AP008212_068 6800001 6910000
AP008212_069 6900001 7010000
AP008212_070 7000001 7110000
AP008212_071 7100001 7210000
AP008212_072 7200001 7310000
AP008212_073 7300001 7410000
AP008212_074 7400001 7510000
AP008212_075 7500001 7610000
AP008212_076 7600001 7710000
AP008212_077 7700001 7810000
AP008212_078 7800001 7910000
AP008212_079 7900001 8010000
AP008212_080 8000001 8110000
AP008212_081 8100001 8210000
AP008212_082 8200001 8310000
AP008212_083 8300001 8410000
AP008212_084 8400001 8510000
AP008212_085 8500001 8610000
AP008212_086 8600001 8710000
AP008212_087 8700001 8810000
AP008212_088 8800001 8910000
AP008212_089 8900001 9010000
AP008212_090 9000001 9110000
AP008212_091 9100001 9210000
AP008212_092 9200001 9310000

AP008212_093 9300001 9410000
AP008212_094 9400001 9510000
AP008212_095 9500001 9610000
AP008212_096 9600001 9710000
AP008212_097 9700001 9810000
AP008212_098 9800001 9910000
AP008212_099 9900001 10010000
AP008212_100 1000001 10110000
AP008212_101 1010001 10210000
AP008212_102 1020001 10310000
AP008212_103 1030001 10410000
AP008212_104 1040001 10510000
AP008212_105 1050001 10610000
AP008212_106 1060001 10710000
AP008212_107 1070001 10810000
AP008212_108 1080001 10910000
AP008212_109 1090001 11010000
AP008212_110 1100001 11110000
AP008212_111 1110001 11210000
AP008212_112 1120001 11310000
AP008212_113 1130001 11410000
AP008212_114 1140001 11510000
AP008212_115 1150001 11610000
AP008212_116 1160001 11710000
AP008212_117 1170001 11810000
AP008212_118 1180001 11910000
AP008212_119 1190001 12010000
AP008212_120 1200001 12110000
AP008212_121 1210001 12210000
AP008212_122 1220001 12310000
AP008212_123 1230001 12410000
AP008212_124 1240001 12510000
AP008212_125 1250001 12610000
AP008212_126 1260001 12710000
AP008212_127 1270001 12810000
AP008212_128 1280001 12910000
AP008212_129 1290001 13010000
AP008212_130 1300001 13110000
AP008212_131 1310001 13210000
AP008212_132 1320001 13310000
AP008212_133 1330001 13410000
AP008212_134 1340001 13510000
AP008212_135 1350001 13610000
AP008212_136 1360001 13710000
AP008212_137 1370001 13810000
AP008212_138 1380001 13910000
AP008212_139 1390001 14010000
AP008212_140 1400001 14110000
AP008212_141 1410001 14210000
AP008212_142 1420001 14310000
AP008212_143 1430001 14410000
AP008212_144 1440001 14510000
AP008212_145 1450001 14610000
AP008212_146 1460001 14710000
AP008212_147 1470001 14810000
AP008212_148 1480001 14910000
AP008212_149 1490001 15010000
AP008212_150 1500001 15110000
AP008212_151 1510001 15210000
AP008212_152 1520001 15310000
AP008212_153 1530001 15410000
AP008212_154 1540001 15510000
AP008212_155 1550001 15610000
AP008212_156 1560001 15710000
AP008212_157 1570001 15810000
AP008212_158 1580001 15910000
AP008212_159 1590001 16010000
AP008212_160 1600001 16110000
AP008212_161 1610001 16210000
AP008212_162 1620001 16310000
AP008212_163 1630001 16410000
AP008212_164 1640001 16510000
AP008212_165 1650001 16610000

AP008212_166 16600001 16710000
AP008212_167 16700001 16810000
AP008212_168 16800001 16910000
AP008212_169 16900001 17010000
AP008212_170 17000001 17110000
AP008212_171 17100001 17210000
AP008212_172 17200001 17310000
AP008212_173 17300001 17410000
AP008212_174 17400001 17510000
AP008212_175 17500001 17610000
AP008212_176 17600001 17710000
AP008212_177 17700001 17810000
AP008212_178 17800001 17910000
AP008212_179 17900001 18010000
AP008212_180 18000001 18110000
AP008212_181 18100001 18210000
AP008212_182 18200001 18310000
AP008212_183 18300001 18410000
AP008212_184 18400001 18510000
AP008212_185 18500001 18610000
AP008212_186 18600001 18710000
AP008212_187 18700001 18810000
AP008212_188 18800001 18910000
AP008212_189 18900001 19010000
AP008212_190 19000001 19110000
AP008212_191 19100001 19210000
AP008212_192 19200001 19310000
AP008212_193 19300001 19410000
AP008212_194 19400001 19510000
AP008212_195 19500001 19610000
AP008212_196 19600001 19710000

Alignment Scores:

Pred. No.: 1,116-18 Length: 110000
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AP008212_240 (1-110000)

QY 8 GlnGluThrThSerGlnAlaLeuAlaValProValMetCysAlaAsnLeuCysGlyPhe 27
DB 109415 AAGGAGACTCGATGCGACGACGACGAGCCCGGATCCTTGATCATATACCGCGCTTC 109474
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerLysCysTyrArgGluThrValMetGln 47
DB 109475 TTTGGCAGTGGCGGTACCATGTAACATGTGCTCAAGTGCACAGAGATGATATGAAG 109534
QY 48 AlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
DB 109535 CAGGAGCAGGCCAAGCTGGCA-----GCCTCTCTATCGACAGCATTTGTCATGTGGT 109588
QY 68 AlaAlaAlaValAlaGlnProProAlaProValHisGluThrLysLeuThrCysGluValGlu 87
DB 109589 GATTCCGGGAGGAGAACCATTTTGTGCTGCACGCTGAAGTACGCTGTTCACATCGAG 109648
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyrGlnGlnAsnLeuValThrProAla 107
DB 109649 GTGAAGACGCTGTGT-----GCCGAGCCTCTCT 109675
QY 108 AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerArgProGluProAsn 127
DB 109676 GAGATTGCTGCTCTCAGTGAAGGGGCTCACGGTGAACCCCAAGGGAGGAGACCAAT 109735
QY 128 ArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysArgCysGlyAsn 147
DB 109736 CCGTCTCTCCACTTGTCCGAAAGGGGTGTCTTACCGGATTCACCTGCCGATGTGGCAAC 109795
QY 148 LeuTyrCysAlaLeuHisArgTyrSerAspLysHisThrCysThrTyrAspTyrLysAla 167
DB 109796 TTGTACTGGCAATGCACCGCTATTCCTGATAGCATGACTGCACGTTGACTATTCGAGAC 109855

QY 168 AlaGlyGlnIleAlaIleAlaLysAlaAsnProLeuValValAlaGluLys 187
DB 109856 GCTCTAGGAGTCTATTGCCAAGGCTATCTCCGTTGGTGAAGGCGGAGA 10991

RESULT 11

AP008212_241

MPCOMMENT

Sequence split into 308 fragments. LOCUS AP008212 Accession A1

Fragment Name	Begin	End
AP008212_000	1	110000
AP008212_001	100001	210000
AP008212_002	200001	310000
AP008212_003	300001	410000
AP008212_004	400001	510000
AP008212_005	500001	610000
AP008212_006	600001	710000
AP008212_007	700001	810000
AP008212_008	800001	910000
AP008212_009	900001	1010000
AP008212_010	1000001	1110000
AP008212_011	1100001	1210000
AP008212_012	1200001	1310000
AP008212_013	1300001	1410000
AP008212_014	1400001	1510000
AP008212_015	1500001	1610000
AP008212_016	1600001	1710000
AP008212_017	1700001	1810000
AP008212_018	1800001	1910000
AP008212_019	1900001	2010000
AP008212_020	2000001	2110000
AP008212_021	2100001	2210000
AP008212_022	2200001	2310000
AP008212_023	2300001	2410000
AP008212_024	2400001	2510000
AP008212_025	2500001	2610000
AP008212_026	2600001	2710000
AP008212_027	2700001	2810000
AP008212_028	2800001	2910000
AP008212_029	2900001	3010000
AP008212_030	3000001	3110000
AP008212_031	3100001	3210000
AP008212_032	3200001	3310000
AP008212_033	3300001	3410000
AP008212_034	3400001	3510000
AP008212_035	3500001	3610000
AP008212_036	3600001	3710000
AP008212_037	3700001	3810000
AP008212_038	3800001	3910000
AP008212_039	3900001	4010000
AP008212_040	4000001	4110000
AP008212_041	4100001	4210000
AP008212_042	4200001	4310000
AP008212_043	4300001	4410000
AP008212_044	4400001	4510000
AP008212_045	4500001	4610000
AP008212_046	4600001	4710000
AP008212_047	4700001	4810000
AP008212_048	4800001	4910000
AP008212_049	4900001	5010000
AP008212_050	5000001	5110000
AP008212_051	5100001	5210000
AP008212_052	5200001	5310000
AP008212_053	5300001	5410000
AP008212_054	5400001	5510000
AP008212_055	5500001	5610000
AP008212_056	5600001	5710000
AP008212_057	5700001	5810000
AP008212_058	5800001	5910000
AP008212_059	5900001	6010000
AP008212_060	6000001	6110000
AP008212_061	6100001	6210000
AP008212_062	6200001	6310000

AP008212_063 6300001 6410000
AP008212_064 6400001 6510000
AP008212_065 6500001 6610000
AP008212_066 6600001 6710000
AP008212_067 6700001 6810000
AP008212_068 6800001 6910000
AP008212_069 6900001 7010000
AP008212_070 7000001 7110000
AP008212_071 7100001 7210000
AP008212_072 7200001 7310000
AP008212_073 7300001 7410000
AP008212_074 7400001 7510000
AP008212_075 7500001 7610000
AP008212_076 7600001 7710000
AP008212_077 7700001 7810000
AP008212_078 7800001 7910000
AP008212_079 7900001 8010000
AP008212_080 8000001 8110000
AP008212_081 8100001 8210000
AP008212_082 8200001 8310000
AP008212_083 8300001 8410000
AP008212_084 8400001 8510000
AP008212_085 8500001 8610000
AP008212_086 8600001 8710000
AP008212_087 8700001 8810000
AP008212_088 8800001 8910000
AP008212_089 8900001 9010000
AP008212_090 9000001 9110000
AP008212_091 9100001 9210000
AP008212_092 9200001 9310000
AP008212_093 9300001 9410000
AP008212_094 9400001 9510000
AP008212_095 9500001 9610000
AP008212_096 9600001 9710000
AP008212_097 9700001 9810000
AP008212_098 9800001 9910000
AP008212_099 9900001 10010000
AP008212_100 10000001 10110000
AP008212_101 10100001 10210000
AP008212_102 10200001 10310000
AP008212_103 10300001 10410000
AP008212_104 10400001 10510000
AP008212_105 10500001 10610000
AP008212_106 10600001 10710000
AP008212_107 10700001 10810000
AP008212_108 10800001 10910000
AP008212_109 10900001 11010000
AP008212_110 11000001 11110000
AP008212_111 11100001 11210000
AP008212_112 11200001 11310000
AP008212_113 11300001 11410000
AP008212_114 11400001 11510000
AP008212_115 11500001 11610000
AP008212_116 11600001 11710000
AP008212_117 11700001 11810000
AP008212_118 11800001 11910000
AP008212_119 11900001 12010000
AP008212_120 12000001 12110000
AP008212_121 12100001 12210000
AP008212_122 12200001 12310000
AP008212_123 12300001 12410000
AP008212_124 12400001 12510000
AP008212_125 12500001 12610000
AP008212_126 12600001 12710000
AP008212_127 12700001 12810000
AP008212_128 12800001 12910000
AP008212_129 12900001 13010000
AP008212_130 13000001 13110000
AP008212_131 13100001 13210000
AP008212_132 13200001 13310000
AP008212_133 13300001 13410000
AP008212_134 13400001 13510000
AP008212_135 13500001 13610000

AP008212_136 13600001 13710000
AP008212_137 13700001 13810000
AP008212_138 13800001 13910000
AP008212_139 13900001 14010000
AP008212_140 14000001 14110000
AP008212_141 14100001 14210000
AP008212_142 14200001 14310000
AP008212_143 14300001 14410000
AP008212_144 14400001 14510000
AP008212_145 14500001 14610000
AP008212_146 14600001 14710000
AP008212_147 14700001 14810000
AP008212_148 14800001 14910000
AP008212_149 14900001 15010000
AP008212_150 15000001 15110000
AP008212_151 15100001 15210000
AP008212_152 15200001 15310000
AP008212_153 15300001 15410000
AP008212_154 15400001 15510000
AP008212_155 15500001 15610000
AP008212_156 15600001 15710000
AP008212_157 15700001 15810000
AP008212_158 15800001 15910000
AP008212_159 15900001 16010000
AP008212_160 16000001 16110000
AP008212_161 16100001 16210000
AP008212_162 16200001 16310000
AP008212_163 16300001 16410000
AP008212_164 16400001 16510000
AP008212_165 16500001 16610000
AP008212_166 16600001 16710000
AP008212_167 16700001 16810000
AP008212_168 16800001 16910000
AP008212_169 16900001 17010000
AP008212_170 17000001 17110000
AP008212_171 17100001 17210000
AP008212_172 17200001 17310000
AP008212_173 17300001 17410000
AP008212_174 17400001 17510000
AP008212_175 17500001 17610000
AP008212_176 17600001 17710000
AP008212_177 17700001 17810000
AP008212_178 17800001 17910000
AP008212_179 17900001 18010000
AP008212_180 18000001 18110000
AP008212_181 18100001 18210000
AP008212_182 18200001 18310000
AP008212_183 18300001 18410000
AP008212_184 18400001 18510000
AP008212_185 18500001 18610000
AP008212_186 18600001 18710000
AP008212_187 18700001 18810000
AP008212_188 18800001 18910000
AP008212_189 18900001 19010000
AP008212_190 19000001 19110000
AP008212_191 19100001 19210000
AP008212_192 19200001 19310000
AP008212_193 19300001 19410000
AP008212_194 19400001 19510000
AP008212_195 19500001 19610000
AP008212_196 19600001 19710000

Alignment Scores:

Pred. No.: 1,11e-18 Length: 110000
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AP008212_241 (1-110000)

/gene="P0429G06.3"
/note="hypothetical ORF
predicted by Genemark.hmm
this category is not included in IRGSP standard"
10066. .12105
/gene="P0429G06.4"
join(10066. .10645,11184. .12105)
/gene="P0429G06.4"
/note="supported by full-length cDNA(s): AK072708"
join(10115. .10645,11184. .11930)
/gene="P0429G06.4"
/note="contains full-length cDNA(s): AK072708"
/codon_start=1
/product="putative polygalacturonase"
/protein_id="BAD3507.1"
/db_xref="GI:51090902"
/translation="MAFRNHDAAVLPFFELMTVTYANAHGSKKPEETQGVGAA
AAVAGPGFTDITKLGAONGRADSTGAVMAAMRSACAGAKITLIPKDGPMGTAME
LRGPGCAVTIQLDGNLGSNDLSKYPGKKNMVEVHRVDFVSGKGLDGGPGV
MSKSCAKVNNCKLPTLVNTVNDVSGITLNAKFRHNTYRCDDIKISGVTIN
APGDSPTNDGIMGDSKTTIAATTIGVDDICTSIGPRTDVNITGVTCGPHGISIG
SLGRYKDERDVAIVTRCVLKKTNGLKISYEDSVPTVSKSYDGVVMDHVDNP
IITDKYCPNSICTSGDSKVSVDVTFRNTITGSSNTPAVVOLLCSGKLPCSGVAMQD
VRVLGSGDKTTAVCDHALGKSTGLKEAL"
complement(join(12464. .12713,12876. .12922))
/gene="P0429G06.5"
complement(join(12464. .12713,12876. .12922))
/gene="P0429G06.5"
/note="P0429G06.5"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
13545. .14247
/gene="P0429G06.6"
13545. .14247
/gene="P0429G06.6"
/gene="P0429G06.6"
/note="P0429G06.6"
/note="probably inactive due to including frameshift(s) in
CDS
pseudogene, polypeptide"
/pseudo
14307. .15085
/gene="P0429G06.7"
14307. .15085
/gene="P0429G06.7"
/note="P0429G06.7"
/note="probably inactive due to including frameshift(s) in
CDS
pseudogene, cytidine deaminase"
/pseudo
complement(16178. .20667)
/gene="P0429G06.8"
complement(join(16178. .16271,16514. .16557,16647. .16815,
16919. .17106,17265. .17327,17685. .17777,19530. .19625,
19717. .19760,19877. .19968,20390. .20667))
/gene="P0429G06.8"
/note="start and end point are not identified"
complement(join(16178. .16271,16514. .16557,16647. .16815,
16919. .17106,17265. .17327,17685. .17777,19530. .19625,
19717. .19760,19877. .19968,20390. .20667))
/gene="P0429G06.8"
/codon_start=1
/product="putative nucleosome assembly protein 1"
/protein_id="BAD3508.1"
/db_xref="GI:51090903"
/translation="WRRPQDLRLGRDLLEMSMNEVEDDEYKPLGRVESMDTEG
EVALRCHRGARGAIVASAVRGSGNTGAGQRMWGAQDRASGADPPGSPSCSKC
RLITPESITATISHRSTNGIELSEBDAIIVETLKXKLQALAQHVDVLESLSVR
KRVLVMEIQSHDELVEKFEPEKALAEAKYGLYPERKGVDPFNLAMKNET
LAELIHSDEBALKYLDIKWCRIDDPGKFEFFPTNPFKNQVLTXYHMTDED
EPILBKALGTEIEHHPGCLTOVLITKSSSTPRTKTECCSFNFPSPQVDDDD
AKIDNTRAEIQNMERDYLAEAVQBDYDASVDDSEEDNDNDESDERA"
complement(join(22647. .22797,22935. .23077))
/gene="P0429G06.9"
complement(join(22647. .22797,22935. .23077))
/gene="P0429G06.9"

/note="hypothetical ORF
predicted by GENESH
this category is not included in IRGSP sta
complement(25897. .31946)
/gene="P0429G06.10"
complement(join(25897. .25991,26101. .261
26685. .26810,26893. .26970,27059. .27247,
28478. .28591,29269. .29454,29538. .29663,
30031. .30131,30226. .30303,30400. .30645,
/gene="P0429G06.10"
/note="start and end point are not identifi
complement(join(25897. .25991,26101. .2621
26685. .26810,26893. .26970,27059. .27247,
28478. .28591,29269. .29454,29538. .29663,
30031. .30131,30226. .30303,30400. .30645,
/gene="P0429G06.10"
/note="contains EST(s):
AU032823(S14678),D48468(S14678),C97402(C60

Alignment Scores:
Pred. No.: 1,65e-18 Length: 158321
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AP003626 (1-158321)

QY 8 GlnGluThrThrSerGlnAlaProGlnGlyProValMetCysGlyValAsnI. 27
Db 119108 AAGGAGACTGATGCGACAGCCAGCAAGGCCGATCCCTTCATCAATAA. 11916
QY 28 PheGlySerGlnAlaThrMetGlyLeuGlySerGlyCysGlyArgGlu. 47
Db 119168 TTTGGCACTGGCGGCTACCATGAACATGCTCTCAAGTCCCAAGAG. 11922
QY 48 AlaIysMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAla. 67
Db 119228 CAGGAGCAGGCCAAGCTTGCA-----GCCCTCTATGACAGCATTC. 11926
QY 68 AlaIaIaIaValGlnProProAlaProValHisGluThrIleuThrC. 87
Db 119282 GATTCGGGAGGAACCAATATTGCTGCTGACCGCTGAAGTAGCTTTC. 11934
QY 88 ArgThrMetIleValProHisGlnSerSerSerGlyGlnIleuAspLeu. 107
Db 119342 GTGAAGACGCTTGT-----G-----G-----G-----G. 11936
QY 108 AlaIaIaIaProGlnAlaValIleSerSerIleAlaIaIaProSerArgP. 127
Db 119369 GAGATTGCTGCTGCTGATGAGAGGGGTGACCGGTGAACCCAGGAGG. 11942
QY 128 ArgCysGlySerCysArgGlyArgValGlyLeuThrGlyPheIleCysG. 147
Db 119429 CGGTGCTCCACTTGTGGAAGGGTGTGCTTACCGGATTCACCTGCC. 11946
QY 148 LeuIleCysAlaIleuHisArgGlySerAspIleHisIleThrCysThrTy. 167
Db 119489 TTGACACGCGCAAGCACCCTATTCGATAGCATACACCTCCAGTTTC. 11954
QY 168 AlaGlyGlnGlnAlaIleAlaIleValAsnProLeuValAlaIaGlu. 187
Db 119549 GCTGCTAGGATGCTATTGCAAGGCTAATCCGGTGTGAAGGCGAG. 11960

RESULT 13
AP003711 163568 bp DNA linear
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, 2004
DEFINITION PAC clone: P0417612. 6,
ACCESSION AP003711
VERSION AP003711.3 GI:51090942

SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group)
REFERENCE	Eurytopyta: Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaeae; Oryza.
AUTHORS	1 Saseaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC clone,P0417G12
JOURNAL	2 Published Only in Database (2001)
REFERENCE	2 (bases 1 to 163568)
AUTHORS	Saseaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Direct Submission
JOURNAL	Submitted (30-MAY-2001) Takuji Saseaki, National Institute of Agricultural Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasakim@nias.affrc.go.jp, URL:http://xgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT	On Aug 9, 2004 this sequence version replaced gi:48525320. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/cdb/glimmerm/glmr_form.html), RiceHMM (http://xgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJ. Protein homologs of the coding regions were searched against NCBI Nonredundant Protein database with BLASTp. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. A gene without classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0417G12 clone has an overlap with P0429G06 (DBJ: AP003626) clone at 5' end and with P0691E09 (DBJ: AP004781) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://xgp.dna.affrc.go.jp/Genomeseq.html.
FEATURES	Location/Qualifiers
source	1..163568
gene	/organism="Oryza sativa (japonica cultivar-group)"
gene	/mol_type="genomic DNA"
gene	/culivar="Nipponbare"
gene	/db_xref="taxon:39947"
gene	/chromosome="6"
gene	/clone="P0417G12"
gene	2936..6301
gene	/gene="P0417G12.1"
gene	join(2936..3941,5348..6301)
gene	/note="supported by full-length cDNA (s) : AK071104"
gene	join(3060..3941,5348..5803)
gene	/gene="P0417G12.1"
gene	/note="contains full-length cDNA (s) : AK071104
gene	plant zinc finger proteins"
gene	/codon_start=1
gene	/product="putative WIP1 protein"
gene	/protein_id="BAD35546.1"
CDS	
mRNA	

repeat_unit	/db_xref="GI:51090943"
repeat_unit	/translation="MHTHTODTTLPLPLSLFTVVCILLHPI
repeat_unit	YVSSSMGDDVYTNFLRGYHHHRRSHHLHPPPPPPPPPY/
repeat_unit	PSSPPIREALPLSLTSTSTTHDDHHRRHQDHHHHKQCI
repeat_unit	DDKRESPSAAADQDAEVTVALHILPSPSPSDAAAKNK
repeat_unit	SSRMQVEEGEEDDEAAATAATPLGCASIGIKLTK
repeat_unit	QFSCPCVYKTFNRNNMMHMGHSQYRKPSBSLRGTQPA
repeat_unit	IDHRARPPLKDFRLQTHYRKHGIKPFMCSCGKAFVRD
repeat_unit	GSDFKRRSLKDHIRAFGRGMAAGCIGSPFDLDADDPSSDL
repeat_unit	10423..14032
repeat_unit	/gene="P0417G12.2"
repeat_unit	join(<10423..10459,13208..13706,13930..
repeat_unit	/gene="P0417G12.2"
repeat_unit	/note="start and end point are not identifi
repeat_unit	join(10423..10459,13208..13706,13930..1
repeat_unit	/gene="P0417G12.2"
repeat_unit	/note="predicted by GeneMark.hmm etc."
repeat_unit	/codon_start=1
repeat_unit	/product="hypothetical protein"
repeat_unit	/protein_id="BAD35547.1"
repeat_unit	/db_xref="GI:51090944"
repeat_unit	/translation="MAMHSGINKRGEDVCGTASGSGHGC
repeat_unit	GSGRGGARSGVAGETNARSGRSARSGSWPLATELVETQ
repeat_unit	AAAELETLVAGTAAVSKKRYNRKHSPGQSFSSSTEDY
repeat_unit	TTDADAASPPASSPLSRCLRAPPPASAPRPPPAITDYLR
repeat_unit	complement(16882..23313)
repeat_unit	/gene="P0417G12.3"
repeat_unit	complement(16882..17848)
repeat_unit	/gene="P0417G12.3"
repeat_unit	/note="3' terminal repeat"
repeat_unit	/rpt_type="terminal
repeat_unit	complement(18655..22278)
repeat_unit	/gene="P0417G12.3"
repeat_unit	/note="probably inactive due to including :
repeat_unit	CDS
repeat_unit	pseudogene. Retrovirus-related Pol polyproi
repeat_unit	transposon TNT 1-94"
repeat_unit	/pseudo
repeat_unit	complement(22350..23313)
repeat_unit	/gene="P0417G12.3"
repeat_unit	/note="5' terminal repeat"
repeat_unit	/rpt_type="terminal
repeat_unit	23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	<23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	/note="start and end point are not identifi
repeat_unit	23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	/note="start and end point are not identifi
repeat_unit	23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	/note="predicted by GlimmerM etc."
repeat_unit	/codon_start=1
repeat_unit	/product="hypothetical protein"
repeat_unit	/protein_id="BAD35548.1"
repeat_unit	/db_xref="GI:51090945"
repeat_unit	/translation="MDASRAHSSAVGILNGERRAQQSRRRG
repeat_unit	WKGDVGAGRWPAOQPLFSFTRGN"
repeat_unit	complement(32128..32409)
repeat_unit	/gene="P0417G12.5"
repeat_unit	complement(<32128..>32409)
repeat_unit	/gene="P0417G12.5"
repeat_unit	/note="start and end point are not identifi
repeat_unit	complement(32128..32409)
repeat_unit	/gene="P0417G12.5"
repeat_unit	/note="contains EST(s) : AU166246(E40406)"
repeat_unit	/codon_start=1
repeat_unit	/product="unknown protein"
repeat_unit	/protein_id="BAD35549.1"
repeat_unit	/db_xref="GI:51090946"
repeat_unit	/translation="MEMKKTAVAFVLAIAFAASATAVLASE
repeat_unit	APGAGGAAAAGCAAGAAATSAAVVMAAPPLAIVACS
repeat_unit	33317..33736
repeat_unit	/gene="P0417G12.6"
repeat_unit	<33317..>33736
repeat_unit	gene
repeat_unit	mRNA
repeat_unit	PTAP
repeat_unit	SCSK
repeat_unit	ELAP
repeat_unit	PPP
repeat_unit	CSNN
repeat_unit	AGGG
repeat_unit	IGGT
repeat_unit	CRNN
repeat_unit	YCAC
repeat_unit	AVYH
repeat_unit	DEEA
repeat_unit	HSGA

/gene="P0417G12.6"
/note="start and end point are not identified"
33317..33736
/gene="P0417G12.6"
/note="predicted by Glimmer etc."
/codon_start=1
/product="hypochemical protein"
/protein_id="BAD35550.1"
/db_xref="GI:51090947"
/translation="MPGGRHSPTRRHRTKATLCTRTTADRSVVCRLMRRCVGR
RRAAALHDARAGGAGAGVODGDRPRVGAADPCRMAGVTCVHTSSSSGPRV
GVALGNLSASTCADRCITTCMCOLESKTEAKLN"
complement(join(33966..34170,34708..34781))
/gene="P0417G12.7"
complement(join(33966..34170,34708..34781))
/gene="P0417G12.7"
/note="hypochemical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(35929..36315)
/gene="P0417G12.8"
complement(<35929..>36315)
/gene="P0417G12.8"
/note="start and end point are not identified"
complement(35929..36315)
/gene="P0417G12.8"
/note="similar to Oryza sativa chromosome 3,
SUNBa0013D02.7"
/codon_start=1
/product="hypochemical protein"
/protein_id="BAD35551.1"
/db_xref="GI:51090948"
/translation="MRVQEPHADAFTPLMCHSLTVVALCFGRSRDWCSSATVP
NPRRHQSAGPLPPSPVGNSSPDVEGVADFFGPTLDLMDLMDGDAKOK
GTLPASSSSSYAQRGAERAGRRP"
36448..36768
/gene="P0417G12.9"
36448..36768
/gene="P0417G12.9"
/note="hypochemical ORF
predicted by Glimmer
this category is not included in IRGSP standard"
37025..38599
/gene="P0417G12.10"
join(<37025..37051,37279..37532,38222..38369,
38462..>38599)
/gene="P0417G12.10"
/note="start and end point are not identified"

Alignment Scores:
Pred. No.: 1.71e-18 Length: 163568
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AP003711 (1-163568)
QY 8 GlnGluThrThrsSerGlnAlaProGluGlyProValMetCysValAsnLeuCyArgGlyPhe 27
Db 42873 AAGGAGACTGGATGCCAGACGACGAGCCCGGATCTTGATCATCATTAAGTGGCGCTTC 42932
QY 28 PheGlySerGlnAlaThrMetGlyLeuCySerIlyrArgGluThrValMetGln 47
Db 42933 TTGGCAGTGGCGGATCATGATGTCTCAAGTGCACACAGAGATGATTAAGAAG 42992
QY 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
Db 42993 CAGGAGCAGGCCAAGCTGGCA-----GCTCTCTTATCGACAGCATTTGTCAATGTGTGT 43046
QY 68 AlaAlaAlaValaGlnProProAlaProValIleGluThrIleLeuThrCysGluValaGlu 87

Db 43047 GATTCGGGAGAGAACCAATTATTCGTGTCACGCTGAAGTAGCTGTTC 43106
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyrlGlnGlnAspLeuV 107
Db 43107 GTGAAGACGCTGT-----G 43133
QY 108 AlaAlaAlaProGlnAlaValIlySerSerIleAlaAlaProSerArgP 127
Db 43134 GAGATTCGTGCTGCTCAATGAGGGGCTCACGGTGAACCCCAAGGAGGAG 43195
QY 128 ArgCysGlySerCysArgIlySerArgValGlyLeuThrGlyPheIlyCysAl 147
Db 43194 CGGTCTCCACTTGTCCGAAAGGGTGTGTCTTACCGATTCACACTGTC 43253
QY 148 LeuTyrcysAlaLeuHisArgTyrserysApIlyshIsthrCysThrTyrc 167
Db 43254 TTGTACTCGCAATGACACCGCTAATCCGATAGCATGACTGCCACTTTC 43313
QY 168 AlaGlyGlnGluAlaIleAlaIlyAlaAsnProLeuValAlaGlu 187
Db 43314 GCTGCTAGGAGATGATTTGCGAAGGCTAATCCGATGAGGCGGAG 43372

RESULT 14
AX654090 522 bp DNA linear -2003
LOCUS Sequence 3960 from Patent WO03000898.
DEFINITION AX654090
ACCESSION AX654090
VERSION AX654090.1 GI:29156904
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Buxariyotai; Viridiplantae; Streptophyta; Embryophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; P
Euharidiales; Oryzaceae; Oryza.
REFERENCE
AUTHORS
Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff
Karagiri,F., Quan,Y., Tao,Y., Whitman,S., Xie,Z., Z
TITLE
Plant genes involved in defense against pathogens
JOURNAL
Patent: WO 03000898-A 3960 03-JAN-2003;
Syngetica Participations AG (CH)
FEATURES
source
location/Qualifiers
1..522
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

Alignment Scores:
Pred. No.: 5.56e-21 Length: 522
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47
Query Match: 42.33% Indels: 35
DB: 6 Gaps: 3

US-10-716-089-20 (1-188) x AX654090 (1-522)
QY 8 GlnGluThrThrsSerGlnAlaProGluGlyProValMetCysIlyAsnI 27
Db 10 AAGAGGCTGCTGCTGCACACAGCCGAGGCGCCCATCTTCATCAATAA 69
QY 28 PheGlySerGlnAlaThrMetGlyLeuCySerIlyrArgGluThrValMetGln 47
Db 70 TTGGCAGTGGCGGACCATGATGTCTCAAGTGCACACAGAGATGATTAAGAAG 129
QY 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 60
Db 130 GAGGAGCAGGCCAAGCTTGTCTCTCTTCATTCGATGACATTTGCAATG 189
QY 61 -----AlaGlnAlaThrSerAlaThrAlaAlaAlaValaG 75
Db 190 GGAAGAGCATATTGTTGTGCGCAGGTGTAGCACGGCGGTGGCGGTG- 237

Oy 76 ProValHtsglThrThylLeuThrCysglValGluArgThrMetIleValProHtsgln 95
 Db 238 -----GCTCAGGTCGAGCGGAGAGACCTCGTT----- 264
 Oy 96 SerSerSerTyrgInglInAspLeuValThrProAlaAlaAlaAlaProGlnAlaVallys 115
 Db 265 -----GTGCAGCCTTACCAGTGTCCGCGGACCGGACCGAGGAG 300
 Oy 116 SerSerIleAlaAlaProSerArgProGluProAlaArgCysglYSerCyArgArglyArg 135
 Db 301 GTTGTGTAGTCCCAAGAGTCAAGAGAGCCGAGACCGGTGGCTGACCTGTAGGAAGG 360
 Oy 136 ValGlyLeuThrglyPheLeuCyArgCysglYAsnLeuTyrcyAlaAlaLeuHtIsergTy 155
 Db 361 GTTGGCGTCAAGCGGATTCACAGTCCGCGGTACAGTACTGTGCGGTGACCGCTAC 420
 Oy 156 SerAspLyHtHsrThrCysThrTyArgTyArgAlaAlaGlyGlnGluAlaIleAllys 175
 Db 421 TCCGACAGCATGATGATGCGAGTTCAGCTACCGAGCTCGGTAGGGAGTCCGCAAG 480
 Oy 176 AlaAsnProLeuValValAlaGluLysValVallys 187
 Db 481 GCCAACCCAGTGTGAAGCTGAAGACTCGACCAAG 516
 RESULT 15
 AK061911 991 bp mRNA linear PLN 24-JUL-2003
 LOCUS Oriza sativa (japonica cultivar-group) cDNA clone:001-042-A06, full
 DEFINITION insert sequence.
 ACCESSION AK061911 GI:32971929
 VERSION FLI CDNA; oligo-capping.
 KEYWORDS Oriza sativa (japonica cultivar-group)
 SOURCE Oriza sativa (japonica cultivar-group)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Euphorbiaceae; Oryzae; Oryza.
 REFERENCE 1
 AUTHORS The Rice Full-Length cDNA Consortium, National Institute of
 Agrobiological Sciences Rice Full-Length cDNA Project Team;
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
 Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C.,
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K.,
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
 Naitakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J.,
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J.,
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Otsu, N., Ota, Y.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice
 Science 301 (5631), 376-379 (2003)
 JOURNAL PUBMED 12869764
 REFERENCE 2 (bases 1 to 991)
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoaka, T.,
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, K.,
 Namiki, T., Naitakawa, R., Niihara, J., Nishikawa, K., Nomura, K.,
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,

Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Satoh, H., Saka
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., S
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagam
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka,
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yake
 Yamada, H., Yamamoto, M., Yasunishi, A., Yasaki, J., Yc
 Yoshimura, A.
 TITLE Direct Submission
 JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National In
 Agrobiological Sciences, Department of Molecular Ge
 Laboratory of Gene Expression; 2-1-2 Kannondai, Iba
 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28K full-length cDNA clone:
 rice.
 URL : http://cdna01.dna.affrc.go.jp/cDNA/
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Ya
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K.
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K.,
 Yamamoto, M.
 FAIS Genome Sequencing & Analysis Group: Otsu, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Ko
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, I
 Mizuno, K., Naitakawa, R., Niihara, J., Oka, M., Ryu, R.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q.,
 Yoshimura, A., Matsubara, K., and Murakami, K.
 Genome Exploration Research Group in Riken Genomic
 and Genome Science Laboratory in Riken: Adachi, J.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Han
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hi
 Hiraoaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Mu
 Nakamura, M., Nishikawa, K., Nomura, K., Numasaki, R., Ohn
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shira
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., W
 Yasunishi, A., and Hayashizaki, Y.

K.,
 and
 ter
 ter
 ica

FEATURES
 source
 1..991
 location/Qualifiers
 /organism="Oryza sativa (japonica cultivar
 /mol_type="mRNA"
 /cultiVar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-042-A06"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,116-20 Length: 991
 Score: 416.50 Matches: 82
 Percent Similarity: 57.29% Conservative: 28
 Best Local Similarity: 42.71% Mismatches: 47
 Query Match: 42.33% Indels: 35
 DB: Gaps: 3

US-10-716-089-20 (1-188) x AK061911 (1-991)

Oy 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysLysAsnL
 Db 153 AAGAGGCTGCTGCACGACCGGAGGCGCCATCTTCATCAATA
 Oy 28 PheGlySerGlnAlaThrMetCylLeuCySerSerlyCysTyArgGluT
 Db 213 TTTGGACGCGGCGGACCATGATGCTCCAAAGTCCCAAGAGAGA
 Oy 48 AlalysMetThrAlaLeuAlaGluGlnAlaThrGlnAla-----
 Db 273 GAGGAGCAGGCCAAGCTTGTGCTCTTCATCATGATGATTCATG

27
 212
 47
 272
 60
 332


```
QY 61 -----AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAla 75
Db 333 GGAAGAGCATATATTGTTGCTGTCAGTGTAGACACGCGGTGCGGTG----- 380
QY 76 ProValHLeGluThrLeuThrCysGluValGluArgThrMetIleValProHLeGln 95
Db 381 -----GCTCAGGTGAGGCGAAGACGCTGCTT----- 407
QY 96 SerSerSerTyrGlnGlnAbleuValThrProAlaAlaAlaAlaProGlnAlaValLys 115
Db 408 -----GTGCAGCCTTACCGATGTGCGGGCACACGAGGAG 443
QY 116 SerSerIleAlaAlaProSerArgProGluProAlaAlaArgCysGlySerCysArgLysArg 135
Db 444 GTTGGTGTAGTCCCAAGGTCAAGAGGCGCAACCGGTGCGCTACCTGTAGGAAGAG 503
QY 136 ValGlyLeuThrGlyPheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyr 155
Db 504 GTTGGGCTGACCGGATTCACTGCCGCTCGGTACATGTACTGTGCGTTCACCGCTAC 563
QY 156 SerAspLysHisThrCysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLys 175
Db 564 TCCGACAAGCATGAATGCCAGTTCGACTACCGGACTGCGGATGCGATGCCACGCAAG 623
QY 176 AlaAsnProLeuValValAlaGluLysValValLys 187
Db 624 GCCAACCCAGTGTGAGGCTGAGAACTCGACAAAG 659
```

Search completed: December 8, 2005, 17:40:21
Job time : 3337 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:32:11 ; Search time 472 Seconds
(without alignments)
2654.561 Million cell updates/sec

Title: US-10-716-089-20

Sequence: 1 MATERVSETTSQAEPGPM.....GQEAIAKANPLVAKVKYKF 188

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_epool_p/US10716089/runac_05122005_094343_13131/app_query.fasta_1.327
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-OUTALIGN=200 -THR SCORE=800 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEMPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716089_0CGN_1_1_1096_@runac_05122005_094343_13131 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_21.*
1: geneseqn1980s.*
2: geneseqn1980s.*
3: geneseqn2000s.*
4: geneseqn2001s.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	984	100.0	804	6	ABST70622 Transcrip
2	984	100.0	804	6	AAS95832 CDNA enco
3	957	97.3	1531	6	ABST70614 Transcrip
4	957	97.3	1531	6	AAS95824 Partial c

ID	Score	Query Match	Length	ID	Description
5	425.5	43.2	945	13	ADX51406
6	419.5	42.6	1000	14	ABE66509
7	419.5	42.6	1021	14	ABE66532
8	416.5	42.3	519	11	ACL30227
9	416.5	42.3	522	8	ADA70637
10	416.5	42.3	522	11	ACL27951
11	416.5	42.3	522	13	ADX52834
12	415	42.2	1032	13	ACL34078
13	415	42.2	736	13	ADX47307
14	415	42.2	776	13	ADX53784
15	415	42.2	821	13	ADX47342
16	415	42.2	826	13	ADX46752
17	415	42.2	841	13	ADX50951
18	415	42.2	846	13	ADX50987
19	415	42.2	853	13	ADX54045
20	415	42.2	863	13	ADX52432
21	415	42.2	863	13	ADX53800
22	415	42.2	868	13	ADX33018
23	415	42.2	868	13	ADX54028
24	415	42.2	872	13	ADX51702
25	415	42.2	872	13	ADX53116
26	415	42.2	873	13	ADX53121
27	415	42.2	876	13	ADX52889
28	415	42.2	886	13	ADX12369
29	415	42.2	894	13	ADX48800
30	415	42.2	898	13	ADX61184
31	415	42.2	899	13	ADX50721
32	415	42.2	902	13	ADX52776
33	415	42.2	904	13	ADX47138
34	415	42.2	918	13	ADX61202
35	415	42.2	920	13	ADX47807
36	415	42.2	924	13	ADX60812
37	415	42.2	926	13	ADX52877
38	415	42.2	928	13	ADX52869
39	415	42.2	931	13	ADX53497
40	415	42.2	932	13	ADX51034
41	415	42.2	942	13	ADX53778
42	415	42.2	943	13	ADX56113
43	415	42.2	948	13	ADX10810
44	415	42.2	953	13	ADX51073
45	415	42.2	956	13	ADX53049

ALIGNMENTS

RESULT 1
ABST70622
ID ABST70622 standard; CDNA; 804 BP.
XX
AC ABST70622;
XX
DT 28-NOV-2002 (first entry)
XX
DE Transcription factor stress-related protein (TFSRP), CDNA
XX
KW Transcription factor stress-related protein; TFSP; MYB-1
KW AP2 similar-2 protein; AP2-2; zinc-finger factor-2; ZF-2;
KW ZF-4 protein; ZF-5 protein; CAAT-Box binding factor; proei
KW Sigma factor like protein; SFL-1; transgenic; plant; drosg
KW environmental stress; Physcomitrella patens; high salinity
KW ciliate; fungus; gene; ss.
XX
OS Physcomitrella patens.
XX
PN US2002102695-A1.
XX
PD 01-AUG-2002.
XX
PF 06-APR-2001; 2001US-00828303.
XX
PR 07-APR-2000; 2000US-0196001P.
XX
XX (SILV/) SILVA O D C E.
PA

PA (BOHN/) BOHNERT H J,
PA (THIE/) THIELN N V,
PA (CHEN/) CHEN R.
XX
XX
PI Silva ODE, Bohnert HJ, Thieln NV, Chen R,
XX
XX WPI, 2002-690614/74,
DR P-PSDB, ABG31805.

Alignment Scores:	
pred. No.:	1,986-77
Score:	984.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	804
Matches:	188
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-716-089-20 (1-188) x ABS70622 (1-804)

QY	1	MetaIaThrInuArgValSerGInGInuTrrThSerGlnAlaProGInuG
DB	134	ATGGCCACCGAACCGTGTGTCTAGAGAGACACTCCGAGGCCCTTGAGG
QY	21	CysIysAsnLeuSCysGlyPhePheGlySerGlnAlaThrMetGlyLeuC
DB	194	TCCAGAAACCTTTGGGGCTTCTTCGGCACCGAACGTACCATGGGGTTGT
QY	41	TYrArgGInuThrValMetGlnAlaLysMetThrAlaLeuAlaGlnAla
DB	254	TACCGAGAGACAGTCATGCAAGCAAGAAATGACGGCTTTAGCTAGCAAG
QY	61	AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProV
DB	314	GCTCAGGAGACACTGTGCCACAGCTGTGTGTTCACGCCCCCTCTGTG
QY	81	LysIleuThrCysGlnValGlnuArgThrMetIleValProHISGlnSerS
DB	374	AAAGCTCACATGGGAGGTTGAGAGAACATGATTGTGCCGATCATATTT
QY	101	GlnAspLeuValThrProAlaAlaAlaAlaAlaProGlnAlaValLysSerS
DB	434	CAGAAGCTGGTTAACCCCGCTGAGCTGCCCTCCAGGACAGTAAGTCTCT
QY	121	ProSerAlaProGlnuProAlaAlaArgCysGlySerCysArgLysArgValG
DB	494	CCCTCTAGACCCGAGGCCAATGCATGCCGATCTTCTGACGAGAAAGGTGTGTG
QY	141	PheLysCysArgCysGlyLysAsnLeuTyrCysAlaLeuHisArgTyrSerAla
DB	554	TTTAAAGTGTGGCTGTGGCAACCTTACTGGGCTTTCATATGGATCTGGG
QY	161	CysThrTyrAspTyrCysAlaAlaAlaGlyGlnGlnuAlaIleAlaLysAlaAla
DB	614	TGCACATATGACTACAAAGCCGAGGCGAGGAAGCGATTCCGAAGCTTA
QY	181	ValAlaGlnLysValValLysPhe 188
DB	674	GTGGCCGAGAAAGTTGTCAAGTTT 697

RESULT 2
AAS95832

Transcription factor stress-related protein; AP2 similar-2
 APS-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;
 MYB-1 MYB-1 protein; CANT-box binding factor protein-3; CANT-
 Sigma factor like protein; SFL-1; environmental stress; mo
 68.

PD 18-OCT-2001
XX

PF 06-APR-2001; 2001WO-US011393.
 XX 07-APR-2000; 2000US-0196001P.
 XX (BADI) BASF PLANT SCI GMBH.
 XX Da Costa Silva O, Bohnert HJ, Van Thielén N, Chen R;
 XX MPI 2002-049146/06.
 DR P-PSDB; AAU71800.
 XX New polypeptide, useful for identification of Physcomitrella patens, and
 PT for modulating stress resistance of a plant, comprises an isolated
 PT transcription factor stress-related protein derived from Physcomitrella
 PT patens.
 XX Claim 14; Fig 2D; 129pp; English.
 XX PS
 XX The invention relates to an isolated transcription factor stress-related
 CC protein (TFSRP) (I) selected from an AP2 similar-2 protein (AP2-2), a
 CC zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 protein, ZF-5
 CC protein, a MYB-1 protein (MYB-1), a CAAT-box binding factor protein-3
 CC (CABF-3), a Sigma factor like protein (SFL-1) and its orthologues. Also
 CC described is an isolated TFSRP coding nucleic acid (II) which codes for
 CC (I); and an isolated recombinant expression vector (III) comprising (II),
 CC where expression of (III) in a host cell results in increased tolerance
 CC to environmental stress as compared to a wild type variety of the host
 CC cell. (I) or (II) is useful as markers for specific regions of the TFSRP
 CC genome. (I) or (II) is useful for identification of Physcomitrella patens
 CC and related organisms, mapping of genomes of organisms related to P.
 CC patens, identification studies, determination of P. patens sequence of
 CC interest, evolutionary studies, modulation of TFSRP regions required
 CC for function, modulation of a TFSRP activity, modulation of the
 CC metabolism of one or more cell functions, modulation of the transmembrane
 CC transport of one or more compounds and modulation of stress resistance.
 CC (II) is useful for generating probes and primers for identifying and/or
 CC cloning TFSRP homologues in other cell types and other organisms, as well
 CC as TFSRP homologues from mosses and related species, and for evolutionary
 CC and protein structural studies. AAS95821-AAS95891 represent P. patens
 CC TFSRP coding sequences and PCR primers of the invention
 CC XX
 SQ Sequence 804 BP; 186 A; 214 C; 205 G; 199 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,98e-77 Length: 804
 Score: 984.00 Matches: 188
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-716-089-20 (1-188) x AAS95832 (1-804)
 QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
 DB 134 ATGGCCACCGACGCGTGTCTTCAGAGAGACGACCTCGCAGGCCCTGAGGGTCCAGTTAG 193
 QY 21 CysIlyAsnLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCysSerIlyCys 40
 DB 194 TGCAGAACCTTGGCGGCTTTCGCGCAGCCAAAGTACCATGCGGGTGTGCTCGAAGTGC 253
 QY 41 TyrArgGluThrValMetGlnAlaIlyMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
 DB 254 TACCAGAGAGACGTATGCAACGAGAGATGACGGCTTTAGCTGACCAAGCCACTCAGGCT 313
 QY 61 AlaGlnAlaThrSerSerAlaThrAlaIlyValGlnProProAlaProValHisGlnThr 80
 DB 314 GCTCGGGGACATCTGCCACACCTGCTGCTTCAGCCCCCGCTCTCTGTACATAGAGCC 373
 QY 81 IlyLeuThrCysGluIlyValGluArgThrMetIlyValProHisGlnSerSerSerTyrgln 100
 DB 374 AAGCTCACATGCGAGATTGAGAGAACAAATGATTTGGCCGATCAATCTTCCAGCTATCA 433

QY 101 GlnAepLeuValThrProAlaIlyAlaIlyProGlnAlaValIlySer 120
 DB 434 CAAGACCTGTATCCCGCTGACAGCTGCCCTCGAGGAGTAAATGCTC 493
 QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgIlyValG 140
 DB 494 CCTTATAGACCGAGACCCCAATGATGCGGATCTTCAGAAACGCTTC 553
 QY 141 PheIlyCysArgCysGlyAsnLeuTyrcysAlaLeuHisArgTyrs 160
 DB 554 TTTAAGTGTGCGTGTGGCAACCTCTACGCTTACATCGGTACTCGG 613
 QY 161 CysThrTyraPtyrIlyAlaIlyGlnGluIlyAlaIlyValIlyAla 180
 DB 614 TGCACATATGACTACAAAGCCGACGGGAGGAAAGCAATTCGAAAGCTA 673
 QY 181 ValAlaGluIlyValIlyPhe 188
 DB 674 GTGGCCGAGAGGTTGTCAAGTTT 697
 RESULT 3
 ID ABS70614 standard; cDNA; 1531 BP.
 XX ABS70614;
 AC 28-NOV-2002 (first entry)
 XX DE Transcription factor stress-related protein (TFSRP), parti
 DT Transcription factor stress-related protein (TFSRP), parti
 XX XX
 KW Transcription factor stress-related protein; TFSRP; MYB-1
 KW AP2 similar-2 protein; AP2-2; zinc-finger factor-2; ZF-2;
 KW ZF-4 protein; ZF-5 protein; CAAT-box binding factor; protei
 KW Sigma factor like protein; SFL-1; transgenic; plant; droug
 KW environmental stress; Physcomitrella patens; high salinity
 KW ciliate; fungus; gene; ss.
 OS Physcomitrella patens.
 XX PN US2002102695-A1.
 XX PD 01-AUG-2002.
 XX PF 06-APR-2001; 2001US-00828303.
 XX PR 07-APR-2000; 2000US-0196001P.
 PA (SILV/) SILVA O D C B.
 PA (BOHN/) BOHNERT H J.
 PA (THIE/) THIELEN N V.
 PA (CHEN/) CHEN R.
 XX PI Silva ODE, Bohnert HJ, Thielén NV, Chen R;
 DR MPI; 2002-690614/74.
 PT Novel transcription factor stress-related protein such as
 PT protein, zinc-finger protein, or CAAT-Box binding factor p
 PT useful for modifying stress tolerance of plant.
 PS Example 5; Fig 1D; 104pp; English.
 XX The invention relates to an isolated transcription factor
 CC protein (TFSRP) (I) selected from AP2 similar-2 protein (A
 CC finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein
 CC MYB-1 protein, CAAT-box binding factor protein-3 (CABF-3),
 CC like protein (SFL-1) and their orthologues. Also described
 CC isolated TFSRP coding nucleic acid (II) coding for (I); (I)
 CC plant cell (III) transformed by (II), where the expression
 CC plant cell results in increased tolerance to an environmen
 CC compared to a wild-type variety of the plant cell; (3) a t
 CC (IV) comprising (III); (4) a seed (V) produced by (IV), wh
 CC true breeding for an increased tolerance to environmental

CC compared to a wild-type variety of the plant cell; (5) an agricultural
CC product (VI) produced by (IV) or (V); and (6) an isolated recombinant
CC expression vector (VII) comprising (II), where expression of (VII) in a
CC host cell results in increased tolerance to environmental stress as
CC compared to a wild-type variety of the host cell. (II) is useful for
CC modifying (increasing or decreasing) stress tolerance of a plant
CC (transgenic or not transgenic), by modifying the expression of TFSRP in
CC the plant. The plant is transformed with a promoter that directs
CC expression of the TFSRP. The promoter is tissue specific and is
CC developmentally regulated. TFSRP expression is modified by administration
CC of an antisense molecule that inhibits expression of TFSRP. (VII) is
CC useful for producing a transgenic plant containing (II), where expression
CC of the nucleic acid in the plant results in increased tolerance to
CC environmental stress as compared to a wild-type variety of the plant, by
CC transforming a plant cell with (VII) comprising the nucleic acid,
CC generating from the plant cell a transgenic plant with an increased
CC tolerance to environmental stress as compared to a wild-type variety of
CC the plant. (I) is useful for conferring stress tolerance such as drought,
CC cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful
CC for identifying Physcomitrella patens and related organisms, for mapping
CC of genomes of organisms related to P.patens, for identifying and
CC localising P.patens sequences of interest, for evolutionary studies, for
CC determining TFSRP regions required for function, for modulating TFSRP
CC activity, for modulating metabolism of one or more cell functions, for
CC modulating transmembrane transport of one or more compounds, and for
CC modulating stress resistance. (I) is useful for reducing stress tolerance
CC response plants or more particularly, in the transcription of a protein
CC involved in a stress tolerance response in a P.patens plant. (II) is
CC useful for transforming plants and thus inducing tolerance to stresses
CC such as drought, high salinity and cold, for identifying the presence of
CC P.patens or a related organism in a mixed population of microorganism,
CC serve as marker for specific regions of the genome for mapping the genome
CC and for the functional studies of P.patens proteins. (II) is useful as
CC reference points for mapping the moss genomes, or of genomes of related
CC organisms, for evolutionary and protein structural studies, for
CC generating knockout mutation in the genomes of various organisms such as
CC bacteria, mammalian cells, yeast cells and plant cells, useful for
CC evaluating their ability or capacity to tolerate various stress
CC conditions and the effect on the phenotype and/or genotype of the
CC mutation. (II) is useful as marker for the construction of a genomic map
CC in related mosses. (I) or (II) is useful for generating algae, ciliates,
CC plants, fungi or other microorganisms expressing mutated TFSRP nucleic
CC acid and protein molecules for improving stress tolerance. ABS70611-
CC ABS70681 represent P. patens TFSRP coding sequences and PCR primers of
CC the invention

XX SQ Sequence 1531 BP, 313 A; 399 C; 393 G; 426 T; 0 U; 0 Other;

Alignment Scores:

Score: 1,046-74 Length: 1531
Percent Similarity: 99.47% Matches: 187
Best Local Similarity: 99.47% Conservative: 0
Query Match: 97.26% Mismatches: 1
DB: 6 Indels: 1
Gaps: 0

US-10-716-089-20 (1-188) x ABS70614 (1-1531)

QY 1 MetAlaThrGluValSerGlnGluThrThrSerGlnAlaProGlnGlyProValMet 20
DB 655 ATGGCCACCGAGCGGTGTCTCAGAGACGACCTCGCAGGCCCTGAGGGTCCACTTATG 714
QY 21 CysIlyAsnLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerIyAsn 40
DB 715 TGCAAGAACCTTGGCGCTCTTCGCGACGCCAAGCTACCATGCGGTGTGCTCGAAGTGC 774
QY 41 TyrArgGluThrValMetGlnAlaIlyMetThrAlaLeuAlaGlnGlnAlaThrGlnAla 60
DB 775 TACCGAGACAGCTCATGAGCGG-AAAGTGAAGCGCTTAACTGACAGACCACTAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaProValHisGlnThr 80
DB * 834 GGTCAAGGCGACATCTGCACACAGCTGCTGTTCAGCCGCCGCTCTGTACATGAGACC 893

QY 81 LysLeuThrCysGluValGluArgThrMetIleValProHisGlnSerS 100
DB 894 AAGCTCATGCGAGGTGTGAGAAACATGATGTGTGGCGATCAATCTT 953
QY 101 GlnAspLeuValThrProAlaAlaAlaProGlnAlaValIlySerS 120
DB 954 CAAGACCTGGTATCCCGCGCTGCGAGCTGCCCTCAGGCGAGTGAAGTCTT 1013
QY 121 ProSerArgProGluProAsnArgCysGlySerCyArgIlyAsnValG 140
DB 1014 CCCCTTGACCCCGAGCCCAATCGATGCGGATCTTGACGAGGAACCGTGTG 1073
QY 141 PheIlyCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSerA 160
DB 1074 TTTAAGTGTCCCTGTGGCACTTACTGCGCTTACATCGCTACTCGG 1133
QY 161 CysThrTyrAspTyrIlyValAlaIleGlnGlnAlaIleAlaIlyValAla 180
DB 1134 TGCACATATGACTCAAAAGCCGAGGCGAGGACGATTCGAAAGCTA 1193
QY 181 ValAlaGluIlyValIlyValPhe 188
DB 1194 GTGGCCGAGAAAGTGTTCAGATT 1217

RESULT 4

AAS95824
ID AAS95824 standard; cDNA; 1531 BP.

AC AAS95824;

XX 26-FEB-2002 (first entry)

DE Partial cDNA encoding zinc finger-4 (ZF-4).

XX Transcription factor stress-related protein; AP2 similar-2

KW AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;

KW MYB-1 MYB-1 protein; CAAT-box binding factor protein-3; CAI

KW Sigma factor like protein; SFL-1; environmental stress; mo

KW 85.

OS Physcomitrella patens.

XX MO200177311-A2.

PD 18-OCT-2001.

XX 06-APR-2001; 2001WO-US011393.

XX 07-APR-2000; 2000US-0196001P.

PA (BAD1) BASF PLANT SCI GMBH.

PI Da Costa Silva O, Bohnert HJ, Van Thielén N, Chen R;

XX WPI, 2002-049146/06.

XX New polypeptide, useful for identification of Physcomitrel.

PT For modulating stress resistance of a plant, comprises an

PT transcription factor stress-related protein derived from Pl

XX patens.

XX Example 5; Fig 1D; 129pp; English.

XX The invention relates to an isolated transcription factor f

CC protein (TFSRP) (I) selected from an AP2 similar-2 protein

CC zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 prc

CC protein, a MYB-1 protein (MYB-1), a CAAT-box binding factor

CC (CABF-3), a Sigma factor like protein (SFL-1) and its ortho

CC described is an isolated TFSRP coding nucleic acid (II) wh

CC (I); and an isolated recombinant expression vector (III) c

CC where expression of (III) in a host cell results in increas

CC to environmental stress as compared to a wild type variety

CC cell. (I) or (II) is useful as markers for specific regions of the TFSRP
CC genome. (I) or (II) is useful for identification of phycomitrella patens
CC and related organisms, mapping of genomes of organisms related to P.
CC patens, identification and localization of P. patens sequence of
CC interest, evolutionary studies, determination of TFSRP regions required
CC for function, modulation of a TFSRP activity, modulation of the
CC metabolism of one or more cell functions, modulation of the transmembrane
CC transport of one or more compounds and modulation of stress resistance.
CC (II) is useful for generating probes and primers for identifying and/or
CC cloning TFSRP homologues in other cell types and other organisms, as well
CC as TFSRP homologues from mosses and related species, and for evolutionary
CC and protein structural studies. AAS95821-AAS95891 represent P. patens
CC TFSRP coding sequences and PCR primers of the invention
XX

XX Sequence 1531 BP; 313 A; 399 C; 393 G; 426 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,04e-74	Length:	1531
Score:	957.00	Matches:	187
Percent Similarity:	99.47%	Conservative:	0
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	97.26%	Indels:	1
DB:	6	Gaps:	0

US-10-716-089-20 (1-188) x AAS95824 (1-1531)

QY 1 MetAATHrGluARValSerGlnGluThrSerGlnAlaProGluGlyProValMet 20
DB 655 ATGGCCACCGAGCGTGTCTCAGGAGAGACGACCTCGACGCCCTGAGGGTCAAGTTATG 714
QY 21 CysIyVaenLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerIyCys 40
DB 715 TGCGAGAACCTTGGGCTTCTTGGCACCAAGCTACATGGGTTGTGCTCGAAGTGC 774
QY 41 TyrArgGluThrValMetGlnAlaIyMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGAGACAGTATGACAGCG-AAAGATGACGGCTTTAGTGAAGAACCTCAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProGlnAlaProValHisGluThr 80
DB 834 GCTCAGCGCACATCGCCACACCTGCTGTTACGCCGCCCTCTGTACATGAGACC 893
QY 81 LysLeuThrCysGluValGluArgThrMetLeuAlaProHisGlnSerSerGlyTyrGln 100
DB 894 AAGCTCACATGGAGAGTTGAGAAACAAGATTTGCGCCATCAATCTTCCAGCTATGCA 953
QY 101 GlnAerPleuValThrProAlaAlaAlaAlaProGlnAlaValIySerSerIleAlaAla 120
DB 954 CAAAGACTGGTTACCCCGCTGACGCTGCCCTCAGGCGATGAAGTCTCTATCCGAGCT 1013
QY 121 ProSerArgProGluProAlaAlaIyGlySerCyArgIyValGlyLeuThrGly 140
DB 1014 CCCTCTAAGACCCGAGCCCAATGATGCGGATCTTTCAGAGAACGTTGGATTGACAGGA 1073
QY 141 PheIySerCyArgCyGlyAlaMetLeuIyCysAlaAlaMetIySerAlaIySerIleAla 160
DB 1074 TTTAAGTGTGCTGGCAACCTCTACATGCGCTTACATCGGTACTCGACAAACACACT 1133
QY 161 CysThrTyrAerPtyIyValAlaIyGlnGlnAlaIleAlaIyAlaAerProLeuVal 180
DB 1134 TGCAATATGACTACAAAGCCGACGAGGAGAGAGATTTGCAAAAGCTTATCTTGTTC 1193

QY 181 ValAlaGluIyValValIySerPhe 188

DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 5

ADXS1406 standard; cDNA; 945 BP.

XX ADXS1406;
XX 21-APR-2005 (first entry)

XX DE Plant full length insert polynucleotide seqid 26146.

XX KW plant protectant; plant growth regulant; gene therapy; pla
XX KW recombinant DNA construct; physical array; plant breeding
XX KW cold tolerance; heat tolerance; drought tolerance; herbic
XX KW extreme osmotic condition; pathogen tolerance; pest toler
XX KW growth rate; cell cycle pathway; disease resistance; plant
XX KW galactomannan production; lignin production; plant growth
XX KW yield; plant growth; plant development; seed oil; protein
XX KW protein content; gene; ss.

OS Unidentified.

XX PN US2004034888-A1.

XX PD 19-FEB-2004.

XX PF 28-APR-2003; 2003US-00425114.

XX PR 06-MAY-1999; 99US-00304517.

XX PR 05-NOV-2001; 2001US-00985678.

XX PA (LNUJ/) LIU J.

XX PA (ZHOU/) ZHOU Y.

XX PA (KOVA/) KOVALIC D K.

XX PA (SCREEN/) SCREEN S E.

XX PA (TABAS/) TABASKA J E.

XX PA (CAOY/) CAO Y.

XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

XX DR WPI; 2004-180133/17.

XX PT New recombinant DNA construct, useful for improving plant

XX PT cold, heat, drought, herbicides, extreme osmotic condition

XX PT pests, for conferring increased resistance to plant diseases

XX PT improving yield.

XX PS Claim 1; SEQ ID NO 26146; 15pp; English.

XX CC The invention describes a recombinant DNA construct compri
XX CC polynucleotide consisting of a sequence encoding an amino
XX CC available in electronic form from the US patent office at
XX CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The
XX CC of the invention are also useful in physical arrays of mol
XX CC plant breeding markers. The recombinant DNA construct is u
XX CC improving plant tolerance to cold, heat, drought, herbicid
XX CC osmotic conditions, pathogens or pests, for manipulating g
XX CC plant cells by modification of the cell cycle pathway, for
XX CC increased resistance to plant disease, for producing galec
XX CC lignin or plant growth regulators, for increasing the rate
XX CC recombinant in plants, for improving yield by modificat
XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use
XX CC or by providing improved plant growth and development unde
XX CC stress condition or for modifying seed oil or protein yiel
XX CC content. This sequence represents a plant full length inee
XX CC polynucleotide that can be used in the recombinant DNA con
XX CC invention.

XX SQ Sequence 945 BP; 207 A; 245 C; 256 G; 237 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.82e-28	Length:	945
Score:	425.50	Matches:	83
Percent Similarity:	61.67%	Conservative:	28
Best Local Similarity:	46.11%	Mismatches:	56
Query Match:	43.24%	Indels:	13
DB:	13	Gaps:	2

US-10-716-089-20 (1-188) x ADXS1406 (1-945)

QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysIyVaenL

```
Db 112 AAGGAGACTGATGTCAGGCACTGAGGAGCCATCTTGATCAATCAATGAGGCTTC 171
Oy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIlySerTyraArgGluThrValMetGln 47
Db 172 TTTCGACAGCGAGCTACATGAACTGCTCCAGTCCCAAGAGATGATATGAG 231
Oy 48 AlalysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
Db 232 CAGAGCAGGCGCAAGCTGGCT-----GCTTCCTATTCGACAGATGCTCAATGGCAAC 285
Oy 68 AlalalalavalGlnProProAlaProValHisGluThrIlyLeuThrCysGluValGlu 87
Db 286 GATGCTGTCAATGAAACA-----GTTGTTGCT 312
Oy 88 ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnApleuValThrProAla 107
Db 313 GGCAACACAGTGTGCTGCTGCTCCAAATCGAGTTGCAACAAATGAACTGACGCCCT 372
Oy 108 AlalalalProGlnAlaValIlySerSerIleAlaIleProSerArgProGluProAsn 127
Db 373 GATGTGCTGACCTAGCGAGGAGGCGGCGGTGATCTCCAAAGGAGAGTGGCCGCAAC 432
Oy 128 ArgCysGlySerCysArgIlyeArgValGlyLeuThrGlyPheIlyCysArgCysGlyAsn 147
Db 433 CGGTGACAGCACTTGACAGAAAGAGGTTGACCTTACAGATTCAACTGCCGGTGTGGGAAAC 492
Oy 148 LeuTyrcysAlaLeuHisArgTyserAspIlyHisIthrCysGthrTyraAspTylyeAla 167
Db 493 TTGTACTGTGACCTGACCGCTACTCCCAAGCAGACGAGTGTGACTATCGACT 552
Oy 168 AlaGlyGlnGluAlaIleAlaIlyeAlaAsnProIleuValValAlaGluIlyeValIlyAs 187
Db 553 CTGCTAGAGATGCCATTGCGCAAGGCTTAATCCAGTGTGAAGCGGACAGATGCACAAAG 612
```

RESULT 6
AEB66909 standard; DNA; 1000 BP.

AC AEB66909;

DT 22-SEP-2005 (first entry)

DE Rice genome derived DNA sequence, SEQ ID 2054.

XX transcription; gene regulation; transgenic plant; RNA interference;

KM transformation; antibody; ds.

OS Oryza sp.

XX JP2005185101-A.

XX 14-JUL-2005.

PF 11-DEC-2002; 2002JP-00383870.

PR 30-MAY-2002; 2002JP-00203269.

XX (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.

PA (SEIB-) SEIBUTSUKEN TOKUTEI SANGYO GIJUTSU.

PA (DOKU-) DOKURITSU GYOSEI HOJIN RIKAGAKU KENKYUSH.

PA (KOKU-) ZH KOKUSAI KAGAKU SHIKO ZAIDAN.

PI Kikuchi H, Hayaehizaki Y, Oromo Y, Matsubara K, Murakami K,

PI Kishimoto N, Sato K, Nagata T, Kawakami N, Yazaki J, Ishikawa M;

PI Doi K, Kawai J;

XX WPI, 2005-566181/58.

XX Novel DNA encoding transcription factor, derived from rice plant, useful

PT for obtaining transcriptional-regulatory regions in plant and for

PT producing modified plant.

XX

XX

XX

PS Claim 1; SEQ ID NO 2054; 2928bp; Japanese.

XX The invention relates to a novel DNA sequence encoding a t-

CC factor derived from a plant. The invention further compris-

CC RNA sequences, ribozyme activity RNA, RNAi sequences, a ve-

CC transformed plant cells, antibodies and proteins, all rela-

CC novel plant DNA sequences of the invention. The novel DNA

CC derived from a rice-genome database. The invention further

CC method for determining the transcription regulatory region;

CC genome. The novel DNA is useful for controlling the expres-

CC in a plant and for producing a modified plant with desired

CC characteristics. The plant DNA and method enables the acqu-

CC transcriptional-regulatory regions. This polynucleotide rei-

CC sequence taken from a rice genome clone library for use in

CC Note: This sequence is not shown in the specification. It i-

CC retrieved from a sequence listing in electronic format fro-

CC Patent Office. The invention claims DNA sequences of SEQ ID

CC 28469 and encoded protein sequences of SEQ ID Nos 28470 to

CC however, the sequence listing only provided the DNA sequenc

CC Nos 1 to 3032.

XX

SO Sequence 1000 BP; 234 A; 242 C; 264 G; 260 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,38e-27	Length:	1000
Score:	419.50	Matches:	82
Percent Similarity:	60.00%	Conservative:	26
Best Local Similarity:	45.56%	Mismatches:	59
Query Match:	42.63%	Indels:	13
DB:	14	Gaps:	2

US-10-716-089-20 (1-188) x AEB66909 (1-1000)

```
Oy 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysIlyAsnI 27
Db 160 AAGGAGACTGATGCCAGCAGCCAGAAAGCCCGATCTTGATCAATAT 219
Oy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIlySerTyraArgGlu 47
Db 220 TTGGCACTGGCGCTTACATGAAACATGCTCTCAAGTGCACAAAGACA 279
Oy 48 AlalysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAla 67
Db 280 CAGAGCAGGCGCAAGCTGGCA-----GCTTCCTATTCGACAGATTC 333
Oy 68 AlalalalavalGlnProProAlaProValHisGluThrIlyeThrC 87
Db 334 GATTCGCGAAGAAACCAATTATGCTGTCACCGCTGAAGTACGTTC 393
Oy 88 ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnApleuV 107
Db 394 GTGAAGACGCTTGT-----G 420
Oy 108 AlalalalProGlnAlaValIlySerSerIleAlaIleProSerArgP 127
Db 421 GAGATTGCTGTCTTACAGAGGCGTCAAGGTAAACCCCAAGGAGAGG 480
Oy 128 ArgCysGlySerCysArgIlyeArgValGlyLeuThrGlyPheIlyCys 147
Db 481 CGGTGCTCCACTTGTCCGAAGAGGTTGCTTACCGGATTCAACTGCC 540
Oy 148 LeuTyrcysAlaLeuHisArgTyserAspIlyHisIthrCysGthrTy 167
Db 541 TTGTACTGTGCGAATGCAACCGCTATTCCGATAGCATGATCCAGTTTG 600
Oy 168 AlaGlyGlnGluAlaIleAlaIlyeAlaAsnProIleuValValAlaG 187
Db 601 CTGCTAGAGATGCTATTGCGCAAGCTTAATCCGTTGTTGAAGCGGAGA 660
```

RESULT 7
AEB66932 standard; DNA; 1021 BP.

ID AEB66932

XX

XX

XX

XX

XX

CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

XX Sequence 519 BP; 117 A; 133 C; 172 G; 97 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,14e-27	Length:	519
Score:	416.50	Matches:	82
Percent Similarity:	57.29%	Conservative:	28
Best Local Similarity:	42.71%	Mismatches:	47
Query Match:	42.33%	Indels:	35
DB:	11	Gaps:	3

US-10-716-089-20 (1-188) x ACL30227 (1-519)

```
QY      8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysLysAsnLeuCysGlyPhe 27
      10 AAGGAGGCTGGCTGCAGAGCCGAGGGCCCAATCTTGTGATCAATAACTGGCGCTTC 69
QY      28 PheGlySerGlnAlaThrMetGlyLeuCysSerLysCysTyrArgGluThrValMetGln 47
      70 TTTGGCAGCGCGCGGCACCATGAACTGTGCTCCAAAGTGCACAGAGATATCATGTGAAG 129
QY      48 AlAluMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
      130 GAGGAGCGAGGCCCACTGCTCCTCCCTCCATGCATAGCATTTGCAATGTTGTACCGT 189
QY      61 -----AlaGlnAlaThrSerAlaThrAlaAlaAlaAlaValGlnProProAla 75
      190 GGGAGGAGCATATTGTTGCTGCAGTGTAGACAGCGGCGTGGCGGTG----- 237
QY      76 ProValHisGluThrLysLeuThrCysGluValGluArgThrMetIleValProHisGln 95
      238 -----GCTCAGGTGAGGGCGAGACCGCTCGTT----- 264
QY      96 SerSerSerTyrGlnGlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLys 115
      265 -----GTGCAGCTTACCGATGTGCGGGGACACGACGAGAG 300
QY      116 SerSerIleAlaAlaProSerArgProGluProAsnArgCysGlySerCysArgLysArg 135
      301 GTTGTGTAGTCCCAAGGTCAAGAGAGGCCGAAACCGGTGCGTAACTGTAGAAAGG 360
QY      136 ValGlyLeuThrGlyPheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyr 155
      361 GTTGGGCTGACGGGATTCACATGCGCGTGGTAACTACTGTGCGTGGACCGCTAC 420
QY      156 SerAspLysHisLeuThrCysThrTyrAspTyrLysAlaIleGluGlnAlaIleAlaLys 175
      421 TCCGCAAGCATGATGCCAGTTCACTACCGAGCTACGGGCTAAGGATGCCATCCCAAG 480
QY      176 AlaAsnProLeuValAlaIleGluLysValValLys 187
      481 GCCAACCACGATGTGAAGCTGAGAAAGCTGCACAAAG 516
DB
```

QY

DB

RESULT 9
ADA70637 standard; DNA; 522 BP.

AC ADA70637;

XX 20-NOV-2003 (first entry)

DT Rice gene, SEQ ID 3960.

XX

KW Plant; bacterial infection; fungal infection; viral infecti
XX gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou
XX Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, ;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance
XX pathogenic infection for conferring resistance or tolerance
XX bacterial, fungal or viral infection by determining or dete

XX gene expression.

XX Claim 6; SEQ ID NO 3960; 899pp; English.

XX The present invention relates to a method (M1) for identifi
XX involved in plant resistance or response to pathogenic inf
XX comprises identifying a gene whose expression is significa
XX the incompatible interaction of plant gene expression rela
XX the expression of the gene in an uninfected plant, in a mutant
XX or in a corresponding incompatible or compatible interacti
XX useful for conferring resistance to resistance or tolerance
XX bacterial, fungal or viral infection. The present sequence
XX illustrate the invention.

XX Sequence 522 BP; 119 A; 133 C; 172 G; 98 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,14e-27	Length:	522
Score:	416.50	Matches:	82
Percent Similarity:	57.29%	Conservative:	28
Best Local Similarity:	42.71%	Mismatches:	47
Query Match:	42.33%	Indels:	35
DB:	8	Gaps:	3

US-10-716-089-20 (1-188) x ADA70637 (1-522)

```
QY      8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysLysAsnL 27
      10 AAGGAGGCTGGCTGCAGAGCCGAGGGCCCAATCTTGTGATCAATAA 69
QY      28 PheGlySerGlnAlaThrMetGlyLeuCysSerLysCysTyrArgGluT 47
      70 TTTGGCAGCGCGGGGACCATGACATGTGCTCCAAAGTGCACAAAGAGA 129
QY      48 AlAluMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
      130 GAGGAGCGAGGCCCAAGCTTGTGCTCCTCCCTCCATGCATAGCATTTGCAATG 189
QY      61 -----AlaGlnAlaThrSerAlaThrAlaAlaAlaValG 75
      190 GGGAGGAGCATATTGTTGCTGCAGTGTAGACAGCGGCGGTG----- 237
QY      76 ProValHisGluThrLysLeuThrCysGluValGluArgThrMetIleV 95
      238 -----GCTCAGGTGAGGGCGAGACCGCTCGTT----- 264
QY      96 SerSerSerTyrGlnGlnAspLeuValThrProAlaAlaAlaProG 115
      265 -----GTGCAGCTTACCGATGTGCGGGGACACGACGAGAG 300
DB
```

QY 116 SerSerIleAlaIaIaProSerArgProGluProAsnArgCysGlySerCysArgIlybArg 135
Db 301 GTTGGCTGTAGTCCCAAGGTCAAGAGAGGCCCAACCGGTGGCTTCTGTGAAAGAG 360
QY 136 ValGlyLeuThrGlyPheIlybCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyr 155
Db 361 GTTGGGCTGACGGGATTTCAACTGCCGTGCGTAAACATGTACTGTCCGTTCACCGCTAC 420
QY 156 SerAspIlybHisIleThrCysThrTyrAspTyrIlybAlaIaGlyGlnGluAlaIleAlaIlyb 175
Db 421 TCCGACAAACATGAAATGCGCATGTTGCACTACCGGACGCGGCTAGGATGCCATCCGCAAG 480
QY 176 AlaAsnProLeuValValAlaGlyIlybValIlyb 187
Db 481 GCCAACCCAGTGTGTGAAGGCTGAAAGCTCGACAAAG 516
RESULT 10
ACL27951
ID ACL27951 standard; cDNA; 522 BP.
AC ACL27951;
DT 02-JUN-2005 (first entry)
XX Rice abiotic stress responsive polynucleotide SEQ ID NO:1907.
XX 88; abiotic stress tolerance; transgenic plant; plant; cereal;
XX agriculture.
XX Oryza sativa.
XX WO2003008540-A2.
XX 30-JAN-2003.
XX 21-JUN-2002; 2002WO-US019668.
XX 22-JUN-2001; 2001US-0300112P.
XX 24-AUG-2001; 2001US-0314662P.
XX 26-SEP-2001; 2001US-0325277P.
XX 21-NOV-2001; 2001US-0332132P.
XX (SYN) SYNGENTA PARTICIPATIONS AG.
XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
XX Moughamer T, Provart N, Rieke D, Zhu T,
XX WPI; 2003-248011/24.
XX PT New stress-responsive nucleic acid, useful for altering the
XX PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX PT stress, salt stress or osmotic stress.
XX PS Claim 1, SEQ ID NO 1907; 89pp; English.
XX CC The invention relates to novel abiotic stress responsive polynucleotides
XX CC and polypeptides. Also disclosed are vectors, expression cassettes, host
XX CC cells, and plants containing such polynucleotides. Also disclosed are
XX CC methods for using the polynucleotides and polypeptides to alter the
XX CC responsiveness of a plant to abiotic stress. The invention is useful in
XX CC agriculture. The nucleic acid is useful for determining whether a test
XX CC plant has been exposed to an abiotic stress condition. It is also useful
XX CC for selecting an agent that alters abiotic stress regulated
XX CC polynucleotide expression in a plant cell, and to identify a homolog or
XX CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
XX CC molecule and the polypeptide encoded by it are useful in altering the
XX CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
XX CC stress, osmotic stress or any of their combinations. The present sequence
XX CC is used in the exemplification of the invention
XX SQ Sequence 522 BP; 119 A; 133 C; 172 G; 98 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,14e-27 Length: 522
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47
Query Match: 42.33% Indels: 35
DB: 11 Gaps: 3
US-10-716-089-20 (1-188) x ACL27951 (1-522)
QY 8 GlnGluThrThrSerGlnAlaIaProGluGlyProValMetCysIlybAsnI 27
Db 10 AAGAGAGCTGCTGCTCCAGCAGAGCCGACGCCCAATCCTTGTGATGATTA 69
QY 28 PheGlySerGlnAlaIaThrMetGlyLeuCysSerIlybCysTyrArgGluT 47
Db 70 TTTGGCAGCGCGCGACCATGAAATGATGTCTCCATGTCGACAGCAAGAGAG 129
QY 48 AlaIlybMetThrAlaLeuAlaGluGlnAlaIaThrGlnAlaIa----- 60
Db 130 GAGGACAGAGCCCAAGCTTGTGCTCTCCATCATGATGATGTCATG 189
QY 61 -----AlaGlnAlaThrSerAlaIaThrAlaAlaIaValG 75
Db 190 GGAAGAGACATATTGTTGCTGCTCCAGTGTACACGCGCGCTG 237
QY 76 ProValHisGluThrIlybLeuThrCysGlyValGluArgThrMetIlyb 95
Db 238 -----GCTCAGGTGCAAGGAGACGCTG 264
QY 96 SerSerSerTyrGlnGlnAspLeuValThrProAlaAlaIaIaIaIaProG 115
Db 265 -----GTGCAAGCTTACCGATGTGCGCGGCA 300
QY 116 SerSerIleAlaIaProSerArgProGluProAsnArgCysGlySerC 135
Db 301 GTTGGCTGTAGTCCCAAGGTCAAGAGAGGCCCAACCGGTGCGTCACT 360
QY 136 ValGlyLeuThrGlyPheIlybCysArgCysGlyAsnLeuTyrCysAlaI 155
Db 361 GTTGGGCTGACGGGATTTCAACTGCCGTGCGTAAACATGTACTGTCCGT 420
QY 156 SerAspIlybHisIleThrCysThrTyrAspTyrIlybAlaIaGlyGlnG 175
Db 421 TCCGACAAACATGAAATGCGCATGTTGCACTACCGGACGCGGCTAGGATG 480
QY 176 AlaAsnProLeuValValAlaGlyIlybValIlyb 187
Db 481 GCCAACCCAGTGTGTGAAGGCTGAAAGCTCGACAAAG 516
RESULT 11
ADX52834
ID ADX52834 standard; cDNA; 1032 BP.
XX ADX52834;
XX 21-APR-2005 (first entry)
XX Plant full length insert polynucleotide seqid 27574.
XX plant protectant; plant growth regulant; gene therapy; pla
XX recombinant DNA construct; physical array; plant breeding
XX cold tolerance; heat tolerance; drought tolerance; herbici
XX extreme osmotic condition; pathogen tolerance; pest tolera
XX growth rate; cell cycle pathway; disease resistance;
XX galactomanan production; lignin production; plant growth
XX yield; plant growth; plant development; seed oil; protein
XX protein content; gene; ss.
XX Undifferentiated.
XX OS US2004034888-A1.
XX PN
XX XX

PD	19-FEB-2004.
XX	
PF	28-APR-2003; 2003US-0042514.
XX	
PR	06-MAY-1999; 99US-00304517.
PR	05-NOV-2001; 2001US-00985678.
XX	
PA	(LIU/) LIU J.
PA	(ZHOU/) ZHOU Y.
PA	(KOVA/) KOVALIC D K.
PA	(SCRE/) SCREEN S E.
PA	(TABAA/) TABASKA J E.
PA	(CAOY/) CAO Y.
XX	
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX	
DR	WPI; 2004-180133/17.
XX	
PT	New recombinant DNA construct, useful for improving plant tolerance to
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT	pests, for conferring increased resistance to plant disease, or for
PT	improving yield.

PS Claim 1, SEQ ID NO 27574, 15bp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 1032 BP; 211 A; 278 C; 304 G; 239 T; 0 U; 0 Other;

```

Db      353  GCTGTGGCCCAAGTGTGAGTTGCACCAACATGCAAGTGCAGCAG-----
Qy      108  -----AlaIlaIaIaProGlnIlaIaValIysSerSerIleIlaIaProS
                |||          |||          |||          |||          |||
Db      401  GATGTTGCCGAGCAGCAGGAGGGGTGGCGGCGCGATCTCCAAAGCGG
                |||          |||          |||          |||          |||
Qy      126  ProAnaIaIaGlyIaSerIaArgIaIaArgIaIaGlyIaLeuThrIaIaPheL
                |||          |||          |||          |||          |||
Db      461  CCGAACCGGTGTAGCGCGCTGTGCAGGAAGAAGGGTGTGCGCTCACGGGATTCa
                |||          |||          |||          |||          |||
Qy      146  GlyaenLeuTyTyCyaaIaIaLeuIaIaArgTyTySerAspIaIaHisThrCysT
                |||          |||          |||          |||          |||
Db      521  GGGAACTTGTACTGTGGCGCGCTCCACCGCTACTCCGACaAGACaCAGCTGCA
                |||          |||          |||          |||          |||
Qy      166  LysIaIaIaIaGlyGlnGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa
                |||          |||          |||          |||          |||
Db      581  CGGACGTGCTGCAGGAGGACGCCATTGTCCAGAGGCTAAATCCGTGGTGAAGG
                |||          |||          |||          |||          |||
Qy      186  ValIys 187
                |||          |||
Db      641  GACCAAG 646

```

RESULT 12

ACCE34078 standard: cDNA: 522 BP.

AC ACL34078;

DT 02-JUN-2005 (first entry)

DE Rice abiotic stress response related polynucleotide SEO ID

XX
XX
es: abiotic stresses to tolerance. Transgenic plant: cere

KW agriculture.

Oryza sativa.

PN WO2003008540-A2.

30-JAN-2003

XX 21-JTN-2002: 2002WO-US0196668-
PF

XX
22 TTN-2001 : 2001HS-0300113B

PR 24-AUG-2001; 2001US-0314662P.

PR 21-NOV-2001; 2001US-0332132P.

PA (SYGN) SYNGENTA PARTICIPATING

aa
PI Kreps J, Briggs SP, Cooper B, Gil

PL Moughnamer T, Provart N, Rieke D, and I;
XX

DR WPI; 2003-248011/24.

New stress-responsive nucleic acid, useful for altering the

PT stress, salt stress or osmotic stress.

PS Claim 48; SEQ ID NO 12641; 89pp; English.

molecule and the polypeptide encoded by it are useful in a

CC responsiveness of a plant to an abiotic stress, such as cold stress, salt stress, osmotic stress or any of their combinations. The present sequence is used in the exemplification of the invention

XX Sequence 522 BP; 122 A; 144 C; 168 G; 88 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,55e-27	Length:	522
Score:	415.00	Matches:	85
Percent Similarity:	61.88%	Conservative:	27
Best Local Similarity:	46.96%	Mismatches:	57
Query Match:	42.17%	Indels:	12
DB:	11	Gaps:	4

US-10-716-089-20 (1-188) x ACU34078 (1-522)

```
OY 8 gngljtrhrthrsrglnalaprrogluglyprovalmetcylsyaanleucysglyphe 27
   |||||
DB 10 AAGAGGCTGGGTGCGAGGCCCGGAGGAGCCATCTCTGCATCAATAACTGTGGCTTC 69
   |||||
OY 28 pheglserglnalathmetglyleucyserysctyrrargluthrvalmetgin 47
   |||||
DB 70 TTCGGCAGCGCGGACCATGACATGCTCTCCAGTGCACAGAGATGATTAACGAAG 129
   |||||
OY 48 AlalysmetthrAlaleuaglgluinalathrglnalaglinalathrserslathr 67
   |||||
DB 130 CAGGATCAGGCCAAGCTGCTCTCTCTCTATGACACAGATCTGTAACGGCAGC 183
   |||||
OY 68 Alalalalavalalgnproprolaaprovalhsgluthrlyleuethrcysgluvalglu 87
   |||||
DB 184 GACGGCGATGAGAGCGGCTGTGCTGCGACAGACAGC-----GTAGTA 228
   |||||
OY 88 ArgthrmetilvalprohlsnserSersrTyrglnGlaaspleuValthrProala 107
   |||||
DB 229 GCTGTGCGCCCAAGTTGAGTTGCAACAATGAACGTGCAGCAG-----CCCGCT 276
   |||||
OY 108 ---AlalalalPRoglnalavalalysSerserilAlalalProserarproglupro 126
   |||||
DB 277 GATGTGTCGCGAACCAGCGAGGCGGCGGATCTCCAAAGGGCGGAGGTAGGCGCG 336
   |||||
OY 127 AenatrgCysglYserCysarglYsarYglYleuthrGlyPheleCysarGly 146
   |||||
DB 337 AACCGGTGAGGCGCTCGAGGAAGAGGTTGAGACTTACCGGATTCACCGCGTGGG 396
   |||||
OY 147 AenleuTYrCysAlaleuHlsarglYsersarplyVhlsfthCysrthrTyrrapTYrlys 166
   |||||
DB 397 AACCTGTACTGGCGACTCCACCGCTACTCCGACAAAGCACTGCACAGTTGACATACCGG 456
   |||||
OY 167 AlalaglYglnGlnalatlAlalysAlaenproleuValAlaglulysValval 186
   |||||
DB 457 ACTGCTGCCAGGAGCGCATTCGCAAGGCTAATCCGTGTGTAAGGACAGCAAGCTCGAC 516
   |||||
OY 187 lys 187
   |||||
DB 517 AAG 519
   |||||
RESULT 13
ID ADX47307
XX ADX47307 standard; cDNA; 736 BP.
AC ADX47307;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polynucleotide seqid 22047.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomanan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
```

KW protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/) LIU J.

XX (ZHOU/) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABA/) TABASKA J E.

XX (CAO/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

XX WPI, 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant

XX cold, heat, drought, herbicides, extreme osmotic condition

XX pests, for conferring increased resistance to plant diseases

XX improving yield.

XX Claim 1, SEQ ID NO 22047; 15pp; English.

CC The invention describes a recombinant DNA construct compri
CC polynucleotide consisting of a sequence encoding an amino
CC available in electronic form from the US patent office at
CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yie
CC content. This sequence represents a plant full length inse
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

CC Sequence 736 BP; 158 A; 213 C; 225 G; 140 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.37e-27	Length:	736
Score:	415.00	Matches:	85
Percent Similarity:	61.88%	Conservative:	27
Best Local Similarity:	46.96%	Mismatches:	57
Query Match:	42.17%	Indels:	12
DB:	13	Gaps:	4

US-10-716-089-20 (1-188) x ADX47307 (1-736)

```
OY 8 gngljtrhrthrsrglnalaprrogluglyprovalmetcylsyaanle 27
   |||||
DB 113 AAGAGGCTGGGTGCGAGGCCCGGAGGAGCCATCTCTGCATCAATTA 172
   |||||
OY 28 pheglserglnalathmetglyleucyserysctyrrargluthrvalmetgin 47
   |||||
DB 173 TTCGGCAGCGCGGACCATGACATGCTCTCCAGTGCACAGAGATGATTAACGAAG 232
   |||||
OY 48 AlalysmetthrAlaleuaglgluinalathrglnalaglinalathrserslathr 67
   |||||
DB 233 CAGGATCAGGCCAAGCTGCTCTCTCTCTATGACACAGATCTG 286
   |||||
```

```
QY 68 AAlaAlaValGlnProProAlaProValHisGluThrLysLeuThrCysGluValGlu 87
XX ||| : : : : : |||
DB 287 GAGCGCCGATGAGACCGGTTGCTGCGACCAACAGC-----GTAATA 331
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnAspLeuValThrProAla 107
XX ||| : : : : : |||
DB 332 GGTGTGGCCCAAGTTGAGTTGCAGAAATGAACGTGCAGAG-----CCCGCT 379
QY 108 ---AlaAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluPro 126
XX ||| : : : : : |||
DB 380 GATGTGGCCGAGCCAGAGGAGGGGTGGCGCATCTCCAAAGGGGAGAGGTAGAGCGCG 439
QY 127 AsnArgCysGlySerCysArgIleValGlyLeuThrGlyPheLysCysArgCysGly 146
XX ||| : : : : : |||
DB 440 AACCGGTGACAGCGCTGCGAGAGAGGGTTGAGCTTAACGGATTAACCTCCGGTGTGGG 499
QY 147 AsnLeuTyrcysAlaLeuHisArgTyrsSerAspLysHisThrCysThrTyrsAspTyrllys 166
XX ||| : : : : : |||
DB 500 AACCTGTACTGTGCACTCCACCGCTACTCCAGCAAGCAAGACTGCAAGTTGCATACCGG 559
QY 167 AlaAlaGlyGlnGlnAlaAlaIleAlaLysAlaAsnProLeuValValAlaGluLysValVal 186
XX ||| : : : : : |||
DB 560 ACTGCTGCCAGGAGACCGCATTCGCAAGGCTAATCCGTTGTGAAGCAGACAAAGCTCGAC 619
QY 187 Lys 187
XX |||
DB 620 AAG 622
XX |||

RESULT 14
ADXS3784
ID ADXS3784 standard; cDNA; 776 BP.
XX
AC ADXS3784;
XX
DT 21-APR-2005 (first entry)
XX
XX Plant full length insert polynucleotide seqid 28524.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; 88.
OS Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 98US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX MPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX
```

```
PT improving yield.
XX
XX Claim 1; SEQ ID NO 28524; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising
XX polynucleotide consisting of a sequence encoding an amino acid
XX available in electronic form from the US patent office at
XX ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The
XX of the invention are also useful in physical arrays of molecules
XX plant breeding markers. The recombinant DNA construct is useful
XX improving plant tolerance to cold, heat, drought, herbicide,
XX osmotic conditions, pathogens or pests, for manipulating growth
XX plant cells by modification of the cell cycle pathway, for
XX increased resistance to plant disease, for producing galactomannan
XX lignin or plant growth regulators, for increasing the rate of
XX reproduction in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen and phosphorus use
XX or by providing improved plant growth and development under
XX stress condition or for modifying seed oil or protein yield
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construction.
XX
XX Sequence 776 BP; 162 A; 225 C; 238 G; 151 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2,52e-27 Length: 776
XX Score: 415.00 Matches: 85
XX Percent Similarity: 61.88% Conservative: 27
XX Best Local Similarity: 46.96% Mismatches: 57
XX Query Match: 42.17% Indels: 12
XX DB: 13 Gaps: 4
XX
XX US-10-716-089-20 (1-188) x ADXS3784 (1-776)
QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysLysAsnL 27
XX ||| : : : : : |||
DB 133 AAGAGGGCTGGGTGTCAGAGCCCGAGGAGCAACCATCTCTGCATCAATAT 192
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerLysCysTyraGlu 47
XX ||| : : : : : |||
DB 193 TTCGCAAGCGGGGAGACCATTAACATGTGCTTAAGTGCACAGAGAA 252
QY 48 AlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaGlnAla 67
XX ||| : : : : : |||
DB 253 CAGATCAGGCCAAGCTGCT-----GCTTCTCTATCGACAGCATG 306
QY 68 AlaAlaValGlnProProAlaProValHisGluThrLysLeuThrC 87
XX ||| : : : : : |||
DB 307 GAGCGCGTCATGAGACCGGTTGCTGCGACCAACAG----- 351
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnAspLeu 107
XX ||| : : : : : |||
DB 352 GCTGTGGCCCAAGTCGAGTTGCAGAAATGAACGTGCAGAG----- 399
QY 108 ---AlaAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSer 126
XX ||| : : : : : |||
DB 400 GATGTGGCCGAGCCAGAGGAGGGGTGGCGCATCTCCAAAGGGGAG 459
QY 127 AsnArgCysGlySerCysArgIleValGlyLeuThrGlyPheLysC 146
XX ||| : : : : : |||
DB 460 AACCGGTGACAGCGCTGCGAGAGAGGGTTGAGCTTAACGGATTAAC 519
QY 147 AsnLeuTyrcysAlaLeuHisArgTyrsSerAspLysHisThrCys 166
XX ||| : : : : : |||
DB 520 AACCTGTACTGTGCACTCCACCGCTACTCCAGCAAGCAAGCTCAACT 579
QY 167 AlaAlaGlyGlnGlnAlaAlaIleAlaLysAlaAsnProLeuValVal 186
XX ||| : : : : : |||
DB 580 ACTGCTGCCAGGAGACCGCATTCGCAAGGCTAATCCGTTGTGAAGCGAG 639
QY 187 Lys 187
XX |||
```

Db 640 AAG 642
RESULT 15
ADK47342
ID ADK47342 etandard; cDNA; 821 BP.
XX
XX ADK47342;
AC
XX
DT 21-APR-2005 (first entry)
DE
XX
XX plant full length insert polynucleotide seqid 22082.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
XX
XX Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOV/) ZHOV Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOV/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX WPI: 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1, SEQ ID NO 22082; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX
XX Sequence 821 BP, 171 A; 236 C; 247 G; 167 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2,71e-27 Length: 821

Score: 415.00 Matches: 85
Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 13 Gaps: 4
US-10-716-089-20 (1-188) x ADK47342 (1-821)
QY 8 GlnGlnThrThrSerGlnAlaProGlnGlyProValMetCysValysAsnI 27
Db 113 AAGGAGGCTGGGTGCGGAGGCGCCCGAGGAGGAGCCCATCTCTGTCATATTA 172
QY 28 PheGlySerGlnAlaThrMetClyLeuCySerIysCysIYrYArgGluT 47
Db 173 TTCGGCAGCGCGCGCAGCATGAACATGTCCTCAAGTCCCAAGGAGAG 232
QY 48 AlaIysMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaT 67
Db 233 CAGGATCAGGCCCAAGCTGGT-----GCTTCCTATCGACAGCATCG 286
QY 68 AlaAlaAlaValGlnProProAlaProValHisGlnThrLysLeuThrC 87
Db 287 GACGCCGTCAATGAGCCGTTGCTGCTGCGACAGACAGC----- 331
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnAspLeuV 107
Db 332 GCTGTGCCCAAGTTGAGTTGCAGAACATGAACGTGCAGCAG----- 379
QY 108 --AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerA 126
Db 380 GATGTGCCCGAGCCGAGGAGGCGGTGCGCGCATCTCCAAAGGGGGA 439
QY 127 AsnArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysC 146
Db 440 AACCGGTGCAGCGCCTGCAGAGAGAGGGTTGAGATTACCGGATTCAAGC 499
QY 147 AsnLeuTyrgValAlaLeuHisArgTyrgSerArgLysIsthrCysThrT 166
Db 500 AACTGTACTGCGCACTCCACCGCTACTCCAGCAAGCAGACTGCAGAT 559
QY 167 AlaAlaGlyGlnGlnAlaIleAlaLysAlaAsnProLeuValAlaIAG 186
Db 560 ACTGCTGCCAGGAGCAGCATTCGCAAGGCTAATCCGATGTGAAGGAGCAG 619
QY 187 Lys 187
Db 620 AAG 622

Search completed: December 8, 2005, 16:44:37
Job time : 477 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:33:58 ; Search time 3752 Seconds
(without alignment)
2344.342 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984

Sequence: 1 MATERVSOETTSQAEKGPV.....GGEALAKANPLVAAKVKVF 188

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10716089/runat.05122005.094344.13151/app.query.fasta_1.327
-DB=EST -QFMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-INITS=bits -STRAT=1 -BND=1 -MATRIX=blosum62 -TRANS=human40.cdt -LIST=45
-DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US10716089 @CGN 1.1 8010 @runat.05122005.094344.13151 -NCPU=6 -ICPU=3
-NO MMAP -LARGEDUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	932	94.7	755	3	BJ958521 BJ958521
2	840	85.4	769	3	BJ610653 BJ610653
3	834	84.8	729	3	BJ592532 BJ592532
4	824	84.8	743	3	BJ592791 BJ592791
5	825	83.8	737	3	BJ584946 BJ584946
6	824	83.7	727	3	BJ585055 BJ585055
7	804	81.7	721	3	BJ585014 BJ585014

C	8	786	79.9	702	3	BJ606134	BJ61
C	9	777	79.0	704	3	BJ591761	BJ59
C	10	708	72.0	719	3	BJ941820	BJ94
C	11	687	69.8	664	3	BJ165418	BJ16
C	12	682	69.3	738	3	BJ605573	BJ60
C	13	676	68.7	723	3	BJ598461	BJ59
C	14	668	67.9	736	3	BJ586048	BJ58
C	15	643	65.3	723	3	BJ611574	BJ61
C	16	617	62.7	642	3	BJ602584	BJ60
C	17	602	61.2	672	3	BJ961154	BJ96
C	18	573	58.2	670	3	BJ590372	BJ59
C	19	564	57.3	651	3	BJ961357	BJ96
C	20	564	57.3	652	3	BJ952449	BJ95
C	21	560	56.9	589	3	BJ173073	BJ17
C	22	546	55.5	612	5	BJ052287	BJ05
C	23	529	53.8	644	3	BJ601969	BJ60
C	24	524	53.3	644	3	BJ597686	BJ59
C	25	524	53.3	644	3	BJ599109	BJ59
C	26	516	52.4	651	3	BJ602208	BJ60
C	27	472	48.0	606	3	BJ167125	BJ16
C	28	451	45.8	760	7	CN208865	CN20
C	29	449	45.6	932	7	CV735162	CV73
C	30	448	45.5	708	3	BJ950422	BJ95
C	31	446	45.3	726	3	BJ950225	BJ95
C	32	438.5	44.6	534	7	CK757250	CK75
C	33	436.5	44.4	716	8	DR102626	DR10
C	34	436.5	44.4	740	8	DR387016	DR38
C	35	436.5	44.4	743	8	DR025280	DR02
C	36	436.5	44.4	746	6	CF401326	CF40
C	37	436.5	44.4	798	6	CF395789	CF39
C	38	436.5	44.4	802	6	CF671829	CF67
C	39	436.5	44.4	809	8	DR048233	DR04
C	40	436.5	44.4	818	6	CF402491	CF40
C	41	436.5	44.4	820	7	CO165465	CO16
C	42	436.5	44.4	837	8	DR101171	DR10
C	43	436.5	44.4	865	8	DR014994	DR01
C	44	436.5	44.4	884	8	DR052935	DR05
C	45	436.5	44.4	892	8	DR014741	DR01

ALIGNMENTS

RESULT 1
BJ958521/c 755 bp mRNA linear
LOCUS BJ958521 pphf full-length cDNA library Physcomitrel
patens cDNA clone pphf6m05 3', mRNA sequence.
ACCESSION BJ958521
VERSION BJ958521.1 GI:67698288
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
Bryophyta: Vriidiplantae; Streptophyta; Embryophyta:
Bryophyta: Funariaceae; Funariales; Funariaceae; ph
1 (bases 1 to 755)
Fujita,T., Nishiyama,T., Shin-i,T., Kohara,Y. and H
Physcomitrella patens EST at a stage of the first a
division of protoplasts
Unpublished (2005)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

Protoplasts were inoculated on BCDMG medium for ex
Protoplasts were isolated from the protoplasts, fur
at 25C under continuous light for 2-3 days. The reg
which were rich in cells at a stage during the first
cell division, were collected. Total RNA was extrac
constructing a full-length cDNA library. The datab
clones is available at the PHYSCDBase (http://moos.

FEATURES

SOURCE

Location/Qualifiers
 1..755
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pPhf36m05"
 /tissue_type="regenerated protoplasts (chloronemata)"
 /dev_stage="at the first asymmetric cell division of protoplasts"
 /clone_lib="pPhf full-length cDNA library"
 /note="Protonemata were inoculated on BCDATG medium for every ca. 5 days. Protoplasts were isolated from the protonemata, further incubated at 25C under continuous light for 2-3 days. The regenerated cells, which were rich in cells at a stage during the first asymmetric cell division, were collected. Total RNA was extracted for constructing a full-length cDNA library."

ORIGIN

Alignment Scores:

Pred. No.: 5,716-80 Length: 755
 Score: 932.00 Matches: 177
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 94.72% Indels: 0
 DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BJ958521 (1-755)

12 SerGlnAlaProGluGlyProValMetCysLysAsnLeuCyGlyPhePheGlySerGln 31
 755 TCGAGAGCCCGCTGAGAGGTCACATTAATGCGAAGAACCTTGGCGCTCTTCGCGACCA 696
 32 AlaThrMetGlyLeuCySerSerLeuCySerLeuGlyThrValMetGlnAlaLysMetThr 51
 695 GCTACCATGCGGCTGTGCTCGAAGTCTCAACGAGACATGTCACGAAAGATGACG 636
 52 AlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThrAlaAlaVal 71
 635 GCTTAAAGCTGACCAAGCACTAGAGGTGCTCAGGCGACACTGCGCAACGCTGCTCTGT 576
 72 GlnProProAlaProValHisGluThrLysLeuThrCysGluValAlaGlnThrMetIle 91
 575 CAGCCCCCGCTCCCTGATCATAGACAGCAAGCTGACATGAGAGTGAAGAACATGATT 516
 92 ValProHisGlnSerSerSerLeuGlnAlaPheValThrProAlaAlaAlaPro 111
 515 GTGCCGCAATCATCTTCCAGCTATCAACAAACCTGTTACCCCGCTGCACGCTGCCCT 456
 112 GlnAlaValLysSerSerLeuAlaAlaProSerArgProGluProAsnArgCysGlySer 131
 455 CAGGCAAGTAAGTCTCTATCGAGCTCCCTCTAAGCCCGACCCCAATGATGCGGATCT 396
 132 CysArgLysArgValAlaGlyLeuThrGlyPheLysCysArgCysGlyAsnLeuThrCysAla 151
 395 TGCAGAGAGCGGTGTTGAGATTGACAGATTAAAGTGTGCGTGGCAACCTCACTACGCGT 336
 152 LeuHisArgTyrSerAspLysHisThrCysThrTyrAspTyrLysAlaAlaGlyGlnGlu 171
 335 TTACATCGGTACTCGGACAAACACACTTGACATATGACTACAAAGCCCGACAGGAGAA 276
 172 AlaAlaLeuAlaValAsnProLeuValAlaGluLysValValLysPhe 188
 275 GCGATTGCAAAAGCTAATCTCTGTGCGTGGCCGAGAAAGTTGCAAGTTT 225

RESULT 2

BJ610693/c

769 bp mRNA linear EST 22-OCT-2003

LOCUS

BJ610693 normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata Physcomitrella patens

subsp. patens cDNA clone pPhn12122 3', mRNA sequence.

ACCESSION

BJ610693

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CONTACT

NATIONAL

INSTITUTE

OF GENETICS

1111 YATA,

MISHIMA,

SHIZUOKA

411-8540,

JAPAN

TEL: 81-559-81-6856

FAX: 81-559-81-6855

EMAIL: tshin@genes.nig.ac.jp

A backbone of the vector is basically from pBluescr

that was in vivo excised from a 1-PIC phage vector

2001). 5' end of the cDNA that was digested with Xh

to SalI site of the vector and the 3' end including

ligated to BamHI site of the

vector(5'-gagagagagagatccacacccgtagagagatgttttttttt

used as a 1st 3' primer, and

5'-ggttctgagatgacatcgctgttccagacgagatgactcgagacggnn

5', hairpin primer, giving the following 5' boarder

agcccaatcgcgcgagctcgaaatgctcgagacg). cDNA inser

amplified with conventional T7 and T3 primers. This

cDNA library was generated according to the method

Mishiyama et al. (2003).

Protonemata were blended by the POLYTRON, and then

the BCDATG medium for 13-14 days under the contin

These clones are available from RIKEN Bio Resource

(http://www.brc.riken.go.jp/lab/epd/Eng/index.html)

of Physcomitrella EST clones is available at the PH

(http://mos.nidb.ac.jp).

Location/Qualifiers

1..769

/organism="Physcomitrella patens subsp. pa

/mol_type="mRNA"

/sub_species="patens"

/db_xref="taxon:145481"

/clone="pPhn12122"

/tissue_type="mixture of chloronemata, cau

rhizoid-like protonemata"

/clone_lib="normalized full length cDNA li

chlonemata, caulonemata and rhizoid-like

1"

ORIGIN

Alignment Scores:

Pred. No.: 4,856-71 Length: 769
 Score: 840.00 Matches: 161
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.37% Indels: 0
 DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BJ610693 (1-769)

28 PheGlySerGlnAlaThrMetGlyLeuCySerLysCysTyrArgGluT 47

768 TTCGACAGCCCAAGCTAACATGCGGTTGTGCTCGAAGTCTCAACCGAGAGA 709

48 AlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaT 67

708 GCGAAGATGACGGCTTAAAGCTGAGCAACCATCTGCGGCTCGACGCGA 649

68 AlaAlaAlaValGlnProProAlaProValHisGluThrLysLeuThrC 87

```

Db      648 GCTGCGCTTTCAGCCGCCCGCTCTGTACAGAACCAAGCTCACATCCGAGGTGG 589
Qy      88 ArgThMetIleValIProHISerSerSerTyTGInIAspLeuValThrProAla 107
Db      588 AGAACAAATGATGTGGCGCATCAATCTTCCAGCTATCAACAAAGACTGGTTACCCCGCT 529
Qy      108 AAlaAlaProGlnAlaValIysSerSerIleAlaIAProSerArgProGluProAla 127
Db      528 GAGCGTCCCTTCAGGACGAGAGAGTCTCTATGCGAGCTCCCTTGAGACCCGAGCCAAAT 469
Qy      128 ArgCyGglSerCyArGlybArGValGlyLeuThiGlyPheIysCyArGCGlyAan 147
Db      468 CCAATGGCGATCTTGCAGGAAGCTGTGGATTACAGAGATTTAAGTGTGCTGGGCAAC 409
Qy      148 LeuTYrCyAlaLeuHISarGlySerAspLYSHISerThrCyArThrTYrAspTYrIysAla 167
Db      408 CTCTACTGCGCTTTACATCGGTACTCGACAACACACTTGCACATATGACTACAAAGCC 349
Qy      168 AAlaGlyGlnGluAlaIleAlaIysAlaAsnProLeuValValAlaIaGlybValValIys 187
Db      348 GAGGGGCGAGAAAGCGATTGCGAAAGCTAACTCTTGTGCTGCGCGAGAAAGTTGTCAAG 289
Qy      188 Phe 188
Db      288 TTT 286

RESULT 3
LOCUS   BU592532/c 729 bp mRNA linear EST 22-OCT-2003
DEFINITION BU592532 normalized full length cDNA library, chloronemata, caulonemata and malformed buds Physcomitrella patens subsp. patens cDNA clone pphb6h03 3', mRNA sequence.
ACCESSION BU592532
VERSION   BU592532.1 GI:37834520
KEYWORDS EST.
SOURCE     Physcomitrella patens subsp. patens
ORGANISM   Physcomitrella patens subsp. patens
REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
AUTHORS   Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Haasbe,M.
TITLE      Comparative Genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
PUBMED    12808149
COMMENT    Contact: Tadao Shii-1
            Center For Genetic Resource Information
            National Institute of Genetics
            111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshii@genes.nig.ac.jp
            A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phase vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector(5'- ggaGagAgAgGATCCACCCTGGAGAGTTTCTTTTCTTTTAA-3' was used as a 1st 3' primer, and 5'-ggTTCGAGTCATCGTGTTCACAGACGATGACTCGAGAACCGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCGACGATCGCATTCGCAAGCCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
            Protonemata were blended by the POLYTRON, and then cultivated on the BCAATG medium for 13- 14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database

```

```

FEATURES
source
    of Physcomitrella EST clones is available at the PH
    (http://moses.nibb.ac.jp):
    Location/Qualifiers
        1..729
            /organism="Physcomitrella patens subsp. pa
            /mol_type="mRNA"
            /sub_species="patens"
            /db_xref="taxon:145481"
            /clone="pphb6h03"
            /tissue_type="mixture of chloronemata, cau
            /clone_lib="normalized full length cDNA li
            chloronemata, caulonemata and malformed bu

ORIGIN
Alignment Scores:
Pred. No.: 1,73e-70 Length: 729
Score: 834.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 84.76% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BU592532 (1-729)

Qy      29 GlySerGlnAlaThrMetGlyLeuCySerIysCyArTYrArgGluThr 48
Db      728 GGCAGCCAGCATACCATGGGGTTGTGCTCGAAGTGTCTACCGAGAGCAC 669
Qy      49 LysMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaThr 68
Db      668 AAGATGACGGCTTTAGCTGACGACCAAGCACTCAGGCTGCTCAGGCGACA 609
Qy      69 AAlaAlaValGlnProProAlaProValHISglnThrIysLeuThrCysG 88
Db      608 GCTGCTGTTCAGCCCCCGCTCTCGTATGACGACCAAGCTCACATGCGC 549
Qy      89 ThrMetIleValIProHISerSerSerTyTGInIAspLeuValT 108
Db      548 ACAATGATTGTGCGCGATCAATCTTCCAGCTATCAACAAGACTGGTTA 489
Qy      109 AAlaAlaProGlnAlaValIysSerSerIleAlaIAProSerArgProG 128
Db      488 GCTGCCCTCAGGCGATGAAGTCTTATCGAGCTCCCTCTAGACCCG 429
Qy      129 CyGlySerCyArGlybArGValGlyLeuThiGlyPheIysCyArG 148
Db      428 TCGGATCTTGCAGGAAGGCTGTGGATTGACAGATTTAAGTGGC 369
Qy      149 TYrCyAlaLeuHISarGlySerAspLYSHISerThrCyArThrTYrAsp 168
Db      368 TACTCGCTTTACATCGGTACTCGACAACACACTTGCACATATGACT 309
Qy      169 GlyGlnGlnAlaIleAlaIysAlaAsnProLeuValValAlaGlybVal 188
Db      308 GGGCGAGAGCGATTGCGAAAGCTAACTCTTGTGCTGCGCGAGAAAG 249

RESULT 4
LOCUS   BU592791/c 733 bp mRNA linear EST 22-OCT-2003
DEFINITION BU592791 normalized full length cDNA library, chloronemata and malformed buds Physcomitrella patens cDNA clone pphb6h21 3', mRNA sequence.
ACCESSION BU592791
VERSION   BU592791.1 GI:37834779
KEYWORDS EST.
SOURCE     Physcomitrella patens subsp. patens
ORGANISM   Physcomitrella patens subsp. patens
REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
AUTHORS   Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y.,

```

TITLE Kohara, Y. and Hasebe, M.
Comparative genomics of *Physcomitrella patens* gametophytic
transcriptome and Arabidopsis thaliana: implication for land plant
evolution
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
PUBMED 12808149
COMMENT Contact: Tadasu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS),
that was in vivo excised from a 1-F1C phage vector (Carninci et al.
2001). 5' end of the cDNA that was digested with XhoI was ligated
to SalI site of the vector and the 3' end including polyA tail was
ligated to BamHI site of the
vector(5'- gagAGAGAGAGATCCAAACCTGgAgAgTTTtttttttttttttVA-3' was
used as a 1st 3' primer, and
5'-ggTTCtGAGtCATGcGtTTCcAGAcGcGATGAcTcGAGAcCGNNNN-3' as 2nd
5'-hairpin primer, giving the following 5' boarder sequence,
AGGCCAAATCGCGAGCTCGAATTCGTcGAGAcCG). cDNA insert could be
amplified with conventional T7 and T3 primers. This full-length
cDNA library was generated according to the method described in
Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then cultivated on
the BCDATG medium for 13-14 days under the continuous light.
These clones are available from RIKEN Bio Resource Center
(http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
of *Physcomitrella* EST clones is available at the PHYSCOBASE
(http://moses.nibb.ac.jp).

FEATURES
source
Location/Qualifiers

```
1..733
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphb6h21"
/tissue_type="mixture of chloronemata, caulonemata and
malformed buds"
/clone_lib="normalized full length cDNA library,
chloronemata, caulonemata and malformed buds"
```

ORIGIN
Alignment Scores:

```
Pred. No.: 1.74e-70 Length: 733
Score: 834.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 84.76% Indels: 0
DB: 3 Gaps: 0
```

US-10-716-089-20 (1-188) x BU592791 (1-733)

```

OY 29 GYSGRGlnAlaThrMeGlyLeuCySeRlyScyTYrARgIuThrValMetGlnAla 48
DB 731 GGCACCAAGCAACATCGGGGTTGCTCTGAAGTCTACCGAGACAGATCTGCAAGC 672
OY 49 IysMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThrAla 68
DB 671 AAGATGACGGCTTGTAGTCAGCAAGCACTCAGGCTGCTCAGGACATCTCCACAGCT 612
OY 69 AlaAlaValAlnProProAlaProAlaProAlaHisGluThrTrpValLeuThrCyGlnValGluArg 88
DB 611 GCTGCTGTTCACCCCGCTCTCTGTACTAGACCAACATCAATGACAGGTTAGAGA 552
OY 89 ThrMetIleValProHisGlnSerSerSerSerTYrGlnGlnAlaPLeuValThrProAlaAla 108
DB 551 ACAATGATTGTGCCGATCAATCTTCAGACTATCAACAGAACTGTTACCCCGCTGCA 492
OY 109 AlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluProAlaArg 128
```

```

DB 491 GTCGCCCTCAGCAGCATGAACTCTTATCGCAGCTCCCTCTAGACCCG 432
OY 129 CySGlySerCyArGlyValAlGlyLeuThrGlyPheLysCyArArgC 148
DB 431 TCCGATCTTCAGAGAACGCTGTGATTCAGACAGGATTTAGTGTGCT 372
OY 149 TYrCyAlaLeuHisArgTYrSerAspLysHisThrCyThrTYrAspT 168
DB 371 TACTCGCTTTACATCGGTACTCGGACAAACACTTGCAATATGACTT 312
OY 169 GYlGlnGlnAlaIleAlaLysAlaAsnProLeuValValAlaGluLysV 188
DB 311 GGCAGGAAGCGATTGCGAAAGTAAATCTCTTGTCTGCGCGAGAGAG 252
```

RESULT 5

LOCUS BU584946/c
DEFINITION

BU584946 747 bp mRNA linear
BU584946 normalized full length cDNA library, chlor
caulonemata and malformed buds *Physcomitrella patens*
cDNA clone pphb26k21 3', mRNA sequence.

ACCESSION
BU584946
GI:37826934

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL Kohara, Y. and Hasebe, M.
PUBMED 12808149
COMMENT Contact: Tadasu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

A backbone of the vector is basically from pBluesc
that was in vivo excised from a 1-F1C phage vector
2001). 5' end of the cDNA that was digested with Xh
to SalI site of the vector and the 3' end including
ligated to BamHI site of the
vector(5'- gagAGAGAGAGATCCAAACCTGgAgAgTTTttttttttt
used as a 1st 3' primer, and
5'-ggTTCtGAGtCATGcGtTTCcAGAcGcGATGAcTcGAGAcCGNN
AGGCCAAATCGCGAGCTCGAATTCGTcGAGAcCG). cDNA insert
amplified with conventional T7 and T3 primers. This
cDNA library was generated according to the method
Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then
the BCDATG medium for 13-14 days under the contin
These clones are available from RIKEN Bio Resource
(http://www.brc.riken.go.jp/lab/epd/Eng/index.html).
of *Physcomitrella* EST clones is available at the PH
(http://moses.nibb.ac.jp).

FEATURES
source
Location/Qualifiers

```
1..747
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphb26k21"
/tissue_type="mixture of chloronemata, cau
malformed buds"
/clone_lib="normalized full length cDNA li
chloronemata, caulonemata and malformed bu
```

ORIGIN

Alignment Scores:

Pred. No.:	1,336-69	Length:	747
Score:	825.00	Matches:	1550
Percent Similarity:	98.76%	Conservative:	0
Best Local Similarity:	98.76%	Mismatches:	2
Query Match:	83.84%	Indels:	0
DB:	3	Gaps:	0

US-10-716-089-20 (1-188) X BJ584946 (1-747)

QY		28	PheGlySerGlnAlaThrMetGlyLeuCysserIysCysTyrArgGluThrValMetGln	47
Dd		747	TTCGGCAGCCAAAGMTACCATGGGGGTGTGCTGAAGTGCTCAACGACGACATCATGCACA	688
QY		48	AlAlaIleMethTrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr	67
Dd		687	GCGAAGATGACGGCTTTTAGCTGAGCAACCATTCAAGCTGTTCAGGCCACATCTCCACA	628
QY		68	AlAlaIlaValaGlnProProAlaAProValHisGluThrTrpIleThrCysGluValGlu	87
Dd		627	GCTGCTCTGTTCAAGCCCCCGCTCTGTGTACATGAGCAAGACTCACATGCAAGTTGAG	568
QY		88	ArgThrMetIleValProHisGlnSerSerSerTyrGlnGlnAspIleValThrProAla	107
Dd		567	AGAACATATGATTGTGGCCGATCAATCTTCCAGCTACACAAAGACTGGATTACCCCCTG	508
QY		108	AlaAlaIlaProGlnAlaValIleSesSerIleAlaAlaProSerArgProGluProAsn	127
Dd		507	GCAAGTCCCTCCAGGACAGTAAGTCCCTATTCGAGGCTCCCTTAAGCCGAGCCAAAT	448
QY		128	ArgCysGlySerCysArgIyValArgValGlyLeuThrGlyPheIysCysArgCysGlyAsn	147
Dd		447	CGATGCCGATCTTCAGAGAACCGTGTGATTGACAGAGATTTAAGTGTGCTGTGGCAAC	388
QY		148	LeuTyrCysAlaIleuNH ₂ ArgTyrSerAspIyHisIleThrCysThrTyrAspTrpTyrAla	167
Dd		387	CTTACTAGCGGTTTACATCGGTACTCGGACAAACACTTACACATTAAGACTACAAACCC	328
QY		168	AlaGlyGlnGlnIuaIleAlaIleAlaLysAlaAspProLeuValValAlaGluIyValIlys	187
Dd		327	GCAGGGCAGAGAAGCGATTGCCAAAAGCTATCTCTTGTCGTGGCCGAGAAGTTGTCAAG	268
QY		188	Phe 188	
Dd		267	TTT 265	
RESULT 6	BUS85055/c			
LOCUS	BUS85055	727 bp	mRNA	linear EST 22-OCT-2003
DEFINITION	BUS85055	normalized full length cDNA library,	chromemata,	caulomemta and malformed buds Physcomitrella patens subsp. patens
		cDNA clone pphb26n05 3', mRNA sequence.		
ACCESSION	BUS85055			
VERSION	BUS85055.1	GI:37827043		
KEYWORDS	EST.			
SOURCE	Physcomitrella patens subsp. patens			
ORGANISM	Physcomitrella patens subsp. patens			
REFERENCE				
AUTHORS	Bryophyta; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;			
	Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.			
	1 (bases 1 to 727)			
	Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H.,			
	Uchiyama,I., Kamiya,A., Carninci,F., Hayashizaki,Y., Shinozaki,K.,			
	Kohara,Y. and Hasebe,M.			
	Comparative genomes of Physcomitrella patens gametophytic			
	transcriptome and Arabidopsis thaliana: implication for land plant			
	evolution			
TITLE	Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)			
JOURNAL	12808149			
PUBMED	Contact: Tadao Shin-i			
COMMENT	Center for Genetic Resource Information			
	National Institute of Genetics			

FEATURES

Source

1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshinigenes.nig.ac.jp
A backbone of the vector is basically from pBluescript that was in vivo excised from a 1-FLC phage vector (2001). 5' end of the cDNA that was digested with XbaI to SalI site of the vector and the 3' end including the ligated to BamHI site of the vector (5'- GAAGAGAGAGAGATGATCCACCCCTGAGAGATTTTTTTTTT- used as a 1st 3' primer, and 5'- aggttcctagagctacgctgcttcgacacgagatgacgcgaacacggn 5'-hairpin primer, giving the following 5' boarder AGGCCCAATCGCCGACGACTGCAATTCTCGAAGACCG). cDNA inserts amplified with conventional T7 and T3 primers. The cDNA library was generated according to the method Nishiyama et al. (2003).
Proconemata were blended by the POLYTRON, and then the BCGATG medium for 13-14 days under the continuous These clones are available from RIKEN Bio Resource (<http://www.brc.riken.go.jp/lab/epd/bng/index.html>). of Physcomitrella EST clones is available at the PH (<http://moss.nib.ac.jp/>).

Location/Qualifiers

1. .727

ORIGIN

Alignment Scores:	
Pred. No.:	1,66-69
Score:	824, 159
Percent Similarity:	99, 388
Best Local Similarity:	99, 388
Query Match:	83, 744
EB:	3
Length:	727
Matches:	159
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-10-716-089-20 (1-188) x BJ585055 (1-727)

QY	29	GlySerGlnAlaThrMetGlyLeuCysSerIysCysTrpArgGluThrV	48
Db	726	GGCAGCCAAAGCTACCATGGGGTTTGGCTCGAAAGTGCTACCGAGAGACAG	667
QY	49	IysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrS	68
Db	666	AAGATGACGGCTTTAGTGCAGCAAGCCACTCAGGCTGCTCAGGGCAGAT	607
QY	63	AlaAlaValAlaGlnProProAlaProValHisGluThrIysLeuThrCysG	88
Db	606	GGTGCTGTTCAGCCCCCGCTTCCTGTATCATGACCAAGCTCAATCGC	547
QY	89	ThrMetIleValProHisGlnSerSerSerTyrrGlnIleAspLeuValT	108
Db	546	ACCATGATTTGGCCGCACTCATATCTTCACGCTATCAACAAGCCGTGTTA	487
QY	109	AlaAlaProGlnAlaValIysSerSerIleAlaIleProSerArgProG	128
Db	486	GCTGCCCTCAGGACAGTAAAGTCTCTATGCCACACTCCCTTACACCCG	427
QY	129	CysGlySerCysArgIysArgValAlaIleLeuThrGlyPheIysCysArgC	148
Db	426	TTCGGATCTTGCAAGGAACGTGGTGGATTGACAGATTTAAAGTCCCA	367
QY	149	TyrCysAlaLeuHisArgTyrSerAspIysHisIshrcCysThrTyrArgP	168
Db	366	TACTCGGCTTACATCGGTACTCGGACAAACACACTTGCAACATATGACT	307

Db 582 GAGACCAAGCTCACATGCGAGGTGAGAGAACATGATTGCGCATCAATCTTCCAGC 523

Qy 99 TyrGInGInaPleuValThrProAlaAlaAlaPProGInAlaValIysSerSerIle 118
 |||||
 Db 522 TATCAACAAGACTGTGTAACCCCGCTGAGTCCCTCAGACGAGTAAGCTCTATC 463
 |||||

Qy 119 AlaAlaPProSerArgProGluProAlaAlaPProGInSerArgIysArgValGlyLeu 138
 |||||
 Db 462 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGATCTTGACGAGACGCTTGATG 403
 |||||

Qy 139 ThrGlyPheIysCyArgCySerGlyValLeuIleThrGlyAlaLeuHisArgTySerAspIys 158
 |||||
 Db 402 ACAGATTAAAGTGTGCTGTGGCAACCTCTACCTCGCTTTACATCGGTACTCGACAA 343
 |||||

Qy 159 HisThrCyThrThrTyAspThrTyIleValIleAlaGlyGInGInAlaIleAlaIysAlaAsnPro 178
 |||||
 Db 342 CACACTTCGNCATATGACTACAAAGCCGAGGGCAGAAAGCATTCGAAAGCTATCT 283
 |||||

Qy 179 LeuValValAlaGluIysValValIysPhe 188
 |||||
 Db 282 CTGTGCGTGGCCGAGAGGTTCTCAAGTTT 253
 |||||

RESULT 10
 BU941820 719 bp mRNA linear EST 13-JUN-2005
 LOCUS BU941820 pphf full-length cDNA library Physcomitrella patens subsp.
 DEFINITION patens cDNA clone pphf16106 5', mRNA sequence.
 ACCESSION BU941820 GI:67568996
 VERSION BU941820.1 GI:67568996
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 REFERENCE 1 (bases 1 to 719)
 Fujita,T., Nishiyama,T., Shin-I,T., Kohara,Y. and Hasebe,M.
 Physcomitrella patens EST at a stage of the first asymmetric cell
 division of protoplasts
 JOURNAL Unpublished (2005)
 COMMENT Contact: Tadao Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshtn@genes.nig.ac.jp
 Protonemata were inoculated on BCDATG medium for every ca. 5 days.
 Protonemata were isolated from the protonemata, further incubated
 at 25C under continuous light for 2-3 days. The regenerated cells,
 which were rich in cells at a stage during the first asymmetric
 cell division, were collected. Total RNA was extracted for
 constructing a full-length cDNA library. The database of the EST
 clones is available at the PHYSCODbase (http://moss.nibb.ac.jp).
 location/Qualifiers

FEATURES
 source 1..719
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="Caxon:145481"
 /clone="pph16106"
 /cissue_type="regenerated protoplasts (chloronemata)"
 /dev_stage="at the first asymmetric cell division of
 protoplasts"
 /clone_lib="pphf full-length cDNA library"
 /note="Protonemata were inoculated on BCDATG medium for
 every ca. 5 days. Protoplasts were isolated from the
 protonemata, further incubated at 25C under continuous
 light for 2-3 days. The regenerated cells, which were rich
 in cells at a stage during the first asymmetric cell
 division, were collected. Total RNA was extracted for
 constructing a full-length cDNA library."

ORIGIN

Alignment Scores:

Pred. No.: 2,78e-58 Length: 719
 Score: 708.00 Matches: 132
 Percent Similarity: 83.15% Conservative: 16
 Best Local Similarity: 74.16% Mismatches: 30
 Query Match: 71.95% Indels: 0
 DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BU941820 (1-719)

Qy 11 ThSerGInAlaPProGluGlyProValMetCySlyAsnLeuCySerGlyP 30
 |||||
 Db 3 ACCGCCAGACCCCGAGAGGGCTGTGATGTTAGAAACGCTGGGGTT 62
 |||||

Qy 31 GlnAlaThrMetGlyLeuCySerIysCySerTyArgGluThrValMetG 50
 |||||
 Db 63 CAGGCACTATGGTCTCTGCTCGAAGCTTACAGAGGAGCCGTATC 122
 |||||

Qy 51 ThrAlaLeuAlaGluGInAlaThrGlnAlaAlaGlnAlaThrSerAlaT 70
 |||||
 Db 123 ACCGCGTGGCCGAGCAGCCCACTCAAGCCGCTCAGTGCATCCG 182
 |||||

Qy 71 ValGInProProAlaProValHisGluThrIysLeuThrCySerGlyValG 90
 |||||
 Db 183 GCTCAGCCTCCAGTCTTATGGAAGAGCACAAGAGCTCATTTAGGCCG 242
 |||||

Qy 91 IleValProHisGInSerSerSerTyGInGInAspLeuValThrProA 110
 |||||
 Db 243 ATCAACCCGCTCAATCTTCAAGCCACCATCCAGTTGAGTTGCTCCGG 302
 |||||

Qy 111 ProGInAlaValIysSerSerIleAlaAlaPProSerArgProGluProA 130
 |||||
 Db 303 CCCAGGTTGATGTTGCCCGGTTGCCACACCTCCACACCGGACCA 362
 |||||

Qy 131 SerCyArgIysArgValGlyLeuThrGlyPheIysCySerGlyValG 150
 |||||
 Db 363 TCTTGACAGAAAGCGGTGGATTGACAGATTTCAAGTTCGTCTGTGGCC 422
 |||||

Qy 151 AlaLeuHisArgTySerAspIysHisThrCySerThrTyArgTyIysAla 170
 |||||
 Db 423 GCTTACATCGGATTTCTGACAAACATATGTTGCACCTTATGATTAAGG 482
 |||||

Qy 171 GluAlaIleAlaIysAlaAsnProLeuValAlaIleGluIysValIle 190
 |||||
 Db 483 GAGGCTATTGCAGAAAGCTAATCCTTATGTTGCGCCGAGAAAGTTGTTA 253
 |||||

RESULT 11
 BU165418 664 bp mRNA linear
 LOCUS BU165418 full length cDNA library, chloronemata and
 DEFINITION gametophores Physcomitrella patens subsp. patens cd
 3', mRNA sequence.
 ACCESSION BU165418 GI:18333403
 VERSION BU165418.1 GI:18333403
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Phy
 1 (bases 1 to 664)
 Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishid
 Uchiyama,I., Kamuya,A., Carninci,P., Hayashizaki,Y.,
 Kohara,Y. and Hasebe,M.
 Comparative genomes of Physcomitrella patens gamet
 transcriptome and Arabidopsis thaliana: implication
 evolution
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 PUBMED 12808149
 COMMENT Contact: Tadao Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856


```

QY 37 CysSerLysCySTyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGlnGln 56
DB 670 TGCTGAAGTGTATACAGGAGACCGCATGACAGAGATGACCGCGTGCAGACG 611
QY 57 AlaThrGlnAlaAlaGlnAlaThrSerAlaThrAlaAlaValaGlnProProAlaPro 76
DB 610 GCCACTCAAGCGCGTCAGGTGTGCATCCGCTGCATCCAGTCTCAGCTTCAATTTCT 551
QY 77 ValHisGluThrLysLeuThrCysGluValaGluATGThrMetIleValProHisGlnSer 96
DB 550 ATGGAAGAGACAAAGAGCTATTGAGCGCCGACATGTGATACAAACCGCTCAACT 491
QY 97 SerSerTyrGlnGlnAlaProGlnValaThrProAlaAlaAlaAlaProGlnAlaValLysSer 116
DB 490 TCCACCCACCATCCAGTTCAGGTGTCTCCGTTACAGTACGAGCCCGAGGTGTATTGCT 431
QY 117 SerIleAlaAlaProSerArgProGluProAlaAlaAlaAlaAlaAlaAlaAlaAla 136
DB 430 CCCGTTGCACACCTCCAGACCGGACCAACCGGTGGATCTTGCAGAGAGCGCGTG 371
QY 137 GlysLeuThrGlyPheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSer 156
DB 370 GGATTGACAGGATTTCAAGTGTCTGTGGGCACTTTTGTGCTTTCATCGGTATTCT 311
QY 157 AspLysHisIleThrCysThrTyrAspTyrLysAlaAlaGlnGlnAlaIleAlaLysAla 176
DB 310 GACAAACATGATGGACATTATGATTAACAAGGCTGGACAGAGGCTATTGCACAAAGCT 251
QY 177 AsnProLeuValaValaGlnLysValaValaLysPhe 188
DB 250 AATCTTTAAGTTGTGTGCGGAGAAAGTTGTTAAGTTC 215
RESULT 13
LOCUS BJS98461/c 723 bp mRNA linear EST 22-OCT-2003
DEFINITION BJS98461 normalized full length cDNA library, chloronemata,
caulonemata and rhizoid-like protonemata Physcomitrella patens
subsp. patens cDNA clone pphn23j11 3', mRNA sequence.
ACCESSION BJS98461 GI:37840453
VERSION BJS98461.1
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
AUTHORS Bryopsida; Funariidae; Funariaceae; Funariaceae; Physcomitrella.
1 (bases 1 to 723)
Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H.,
Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
Kohara,Y. and Hasebe,M.
Comparative genomes of Physcomitrella patens gametophytic
crustaceipome and Arabidopsis thaliana: implication for land plant
evolution
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
12808149
Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
A backbone of the vector is basically from pBluescript II (KS),
that was in vivo excised from a 1-FLC phase vector (Carninci et al.
2001). 5' end of the cDNA that was digested with XhoI was ligated
to SalI site of the vector and the 3' end including polyA tail was
ligated to BamHI site of the
vector(5'- gAAGAGAGAGATCCAAACCTGgAGAGTATTTTATTTTATTTTAA-3' was
used as a 1st 3' primer, and
5'-ggTTCGAgATCGTGTTCGAGACGgATGACTGAGAACGNNNN-3' as 2nd
5'-hairpin primer, giving the following 5' boarder sequence,
AGGCGCAATCGCGCGAGCTCAATTGCTCGGAGACG). cDNA instert could be
amplified with conventional T7 and T3 primers. This full-length

```

```

FEATURES
    source
        cdna library was generated according to the method
        Nishiyama et al. (2003).
        the BCDATs medium for 13- 14 days under the continuc
        These clones are available from RIKEN Bio Resource
        (http://www.brc.riken.go.jp/lab/epd/Eng/index.html)
        of Physcomitrella EST clones is available at the PH
        (http://mos.nibb.ac.jp).
    location/Qualifiers
        1..723
        /organism="Physcomitrella patens subsp. pa
        /mol_type="mRNA"
        /sub_species="patens"
        /db_xref="taxon:145481"
        /clone="pphn23j11"
        /tissue_type="mixture of chloronemata, cau
        rhizoid-like protonemata"
        /clone_id="normalized full length cDNA 1:
        chloronemata, caulonemata and rhizoid-like

ALIGNMENT Scores:
Pred. No.: 3.54e-55 Length: 723
Score: 676.00 Matches: 126
Percent Similarity: 83.04% Conservative: 16
Best Local Similarity: 73.68% Mismatches: 29
Query Match: 68.70% Indels: 0
Gaps: 0
DB: 3

US-10-716-089-20 (1-188) x BJS98461 (1-723)
QY 18 ProValMetCysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrM- 37
DB 723 CCTGTGATGTGTAGAAACGTGTGCGGTTCTTTGGACCAAGCCACTA- 664
QY 38 SerLysCysTyrArgGluThrValMetGlnAlaLysMetThrAlaLeu 57
DB 663 TCGAAGGTGTACAGGAGACCGTCATGCAAGCAAGATGACGCGGTG- 604
QY 58 ThrGlnAlaAlaGlnAlaThrSerAlaThrAlaAlaValGlnProP- 77
DB 603 ACTCAAGCGGCTCAGGTGCTGCCATCCGTCATCCAGTGTCAAGCTC- 544
QY 78 HisGluThrLysLeuThrCysGluValaGluArgThrMetIleValProH- 97
DB 543 GAAAGAGACAGAGCTCTTTGAGGCCAGCAGCATGTGATTAACAACGC- 484
QY 98 SerTyrGlnGlnAsnLeuValaThrProAlaAlaAlaAlaProGlnAlaV- 117
DB 483 AGCCACCATCCAGTTGAGGTTGCTCCGTTACAGTAGCCCCCAAGTTG- 424
QY 118 IleAlaAlaProSerArgProGluProAlaAlaAlaAlaAlaAlaV- 137
DB 423 GTTGCCACACCCCTCAGACCGGCAACAAACCGGTGCGATCTTGACAGA- 364
QY 138 LeuThrGlyPheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisA- 157
DB 363 TTGACAGGATTTTCAGTGTGCTGTGGGCACTTTTGTGCTTTTACATC- 304
QY 158 LysHisIleThrCysThrTyrAspTyrLysAlaAlaGlnGlnAlaIleA- 177
DB 303 AAACATAGTGCACATTATGATTAACAAGCTGCTGGACAGAGGCTATTG- 244
QY 178 ProLeuValaValaGlnLysValaValaLysPhe 188
DB 243 CTTTAAAGTTGTGTGCGGAGAAAGTTGTTAAGTTC 211
RESULT 14
LOCUS BJS86048/c 736 bp mRNA linear
DEFINITION BJS86048 normalized full length cDNA library, chlor
caulonemata and malformed buds Physcomitrella pater
cDNA clone pphb29d18 3', mRNA sequence.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 8, 2005, 16:36:03 ; Search time 165 Seconds
(without alignment)
2025.343 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984

Sequence: 1 MATEVSESTTSQABEGPVM.....GGEALAKANPLVAKVKVF 188

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Dgapop 6.0 , Dgapext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cg2n2_1/USPTO_apool_p/US10716089/rn1/05122005_094344_13172/app_query.fasta_1.327
-DB=Issued Patents NA -QFMT=fastp -SUPPLX=p2n.rn1 -MINMATCH=0.1 -LOOPC=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716089 @CGN_1_1_290 @rn1/05122005_094344_13172 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBSTR -NEG_SCORES=0 -WAIT -DISPATCH=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELPO=6 -DELEXT=7

Database :

Issued Patents NA:*

- 1: /cg2n2_6/ptodata/1/lna/1_COMB.seq:*
- 2: /cg2n2_6/ptodata/1/lna/5_COMB.seq:*
- 3: /cg2n2_6/ptodata/1/lna/6A_COMB.seq:*
- 4: /cg2n2_6/ptodata/1/lna/6B_COMB.seq:*
- 5: /cg2n2_6/ptodata/1/lna/H_COMB.seq:*
- 6: /cg2n2_6/ptodata/1/lna/PCUS_COMB.seq:*
- 7: /cg2n2_6/ptodata/1/lna/PP_COMB.seq:*
- 8: /cg2n2_6/ptodata/1/lna/RE_COMB.seq:*
- 9: /cg2n2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	984	100.0	804	3	US-09-828-303-12
2	957	97.3	1531	3	US-09-828-303-4
3	367	37.3	2515	3	US-09-244-805-10
4	365	37.1	2425	3	US-09-949-016-810
5	365	37.1	2684	3	US-09-949-016-3693
6	344.5	35.0	1197	2	US-08-861-269-4
7	344.5	35.0	1197	2	US-09-134-596-4
8	344.5	35.0	1197	3	US-09-293-273-4
9	277	28.2	533	3	US-09-640-211A-1804

10	265	26.9	283	3	US-09-313-294A-7446	Seq
11	261	26.5	281	3	US-09-313-294A-1951	Seq
12	219	22.3	211	3	US-09-313-294A-3856	Seq
13	207	21.0	348	3	US-09-640-211A-1742	Seq
14	193.5	19.7	1281	3	US-09-620-312D-849	Seq
15	193	19.6	14524	3	US-09-949-016-12552	Seq
16	193	19.6	14785	3	US-09-949-016-15435	Seq
17	187	19.0	2089	3	US-10-104-047-106	Seq
18	160	16.3	329	3	US-09-640-211A-1702	Seq
19	159.5	16.2	306	3	US-09-313-294A-5600	Seq
20	159	16.2	181	3	US-09-513-999C-9489	Seq
21	148.5	15.1	277	3	US-09-313-294A-6731	Seq
22	148.5	15.1	291	3	US-09-313-294A-6731	Seq
23	148.5	15.1	307	3	US-09-313-294A-6505	Seq
24	146.5	14.9	389	3	US-09-640-211A-1637	Seq
25	146.5	14.9	362	3	US-09-640-211A-1587	Seq
26	145	14.7	467	3	US-09-640-211A-1667	Seq
27	143.5	14.6	563	3	US-09-640-211A-1657	Seq
28	143	14.5	315	3	US-09-640-211A-1346	Seq
29	143	14.5	322	3	US-09-640-211A-1337	Seq
30	143	14.5	368	3	US-09-640-211A-1379	Seq
31	142	14.4	313	3	US-09-640-211A-1350	Seq
32	142	14.4	317	3	US-09-640-211A-1723	Seq
33	142	14.4	355	3	US-09-640-211A-1657	Seq
34	142	14.4	406	3	US-09-640-211A-1709	Seq
35	137.5	14.0	265	3	US-09-313-294A-3339	Seq
36	134	13.6	473	3	US-09-640-211A-1740	Seq
37	129	13.1	320	3	US-09-640-211A-1371	Seq
38	120.5	12.2	330	3	US-09-640-211A-1290	Seq
39	118	12.0	290	3	US-09-313-294A-5271	Seq
40	117.5	11.9	259	3	US-09-313-294A-161	Seq
41	110	11.2	1863	3	US-09-902-540-4947	Seq
42	110	11.2	30780	3	US-09-902-540-1243	Seq
43	105	10.7	601	3	US-09-949-016-33790	Seq
44	105	10.7	601	3	US-09-949-016-33790	Seq
45	103.5	10.5	2664	2	US-08-942-819-1	Seq

ALIGNMENTS

RESULT 1
US-09-828-303-12
Sequence 12, Application US/09828303
Patent No. 6677504
GENERAL INFORMATION:
APPLICANT: COSTA E SILVA, OSWALDO DA
APPLICANT: BOHNER, HANS J.
APPLICANT: VAN THIELEN, NOCHA
TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROT
TITLE OF INVENTION: METHODS OF USE IN PLANTS
FILE REFERENCE: 16313-0030
CURRENT APPLICATION NUMBER: US/09/828,303
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/196,001
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 804
TYPE: DNA
ORGANISM: Phycocytella patens
US-09-828-303-12

Alignment Scores:

Pred. No.:	6,21e-87	Length:	804
Score:	984.00	Matches:	188
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-716-089-20 (1-188) x US-09-828-303-12 (1-804)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
DB 134 ATGGCCACCGAGCGTGTCTTCAGAGACGACCTCGAGCCCTCGAGGGTCCATTATG 193
QY 21 CysIysAsnLeuCySGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerIysCyS 40
DB 194 TGCAGAAACCTTTGGGGCTTCTTCGGCAGCCCAAGCTACCATGGGGTTGTGCTCGAAGTGC 253
QY 41 TyrArgGluThrValMetGlnAlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 254 TACCCAGAGACAGTCATGCACAGCAAGATGACGGCTTTAGCTGACCAAGCCCTCAGGCT 313
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHisGluThr 80
DB 314 GCTCAGGCGACATCTGCCACAGCTGCTGCTTACGCCCCCGCTCTCTGACATAGACC 373
QY 81 LysLeuThrCySGluValGluArgThrMetIleValProHisGlnSerSerSerTyGln 100
DB 374 AAGCTCACATGCGAGGTTGAGAGAACATGATTTGCCGCATCATCTCCAGCTATCA 433
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValIysSerSerIleAlaAla 120
DB 434 CAAGCCTGGTTACCCCGCTGACAGCTGCCCTCAGGACGTGAACTCTATTCGACACT 493
QY 121 ProSerArgProGluProAsnArgCySGlySerCySArgIysArgValGlyLeuThrGly 140
DB 494 CCTCTAGACCCGAGCCCAATGATCGGATCTTCAGAGAACGCTGTGGATTCAGAGA 553
QY 141 PheIysCySArgCySGlyAsnLeuTyrcySAlaLeuHisArgTyrcySerAspIysHisThr 160
DB 554 TTTAAGTGTGCTGTGGCAACCTCTACGCTTACATCGGTACTCGAGCAACAACACT 613
QY 161 CysThrTyrcyAspTyrcyIysAlaAlaGlyGlnGluAlaIleAlaIysAlaAsnProIleVal 180
DB 614 TGCACATATGACTACAAAGCCGACAGGCGAGAAAGATTCGAAAGCTAATCTCTGTGC 673
QY 181 ValAlaGluIysValValIysPhe 188
DB 674 GTGGCCGAGAGGTTGTCAAGTTT 697

RESULT 2
US-09-828-303-4
; Sequence 4, Application US/09828303
; Patent No. 6677504
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
; TITLE OF INVENTION: METHODS OF USE IN PLANTS
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Physcomitrella patens
US-09-828-303-4

Alignment Scores:
Pred. No.: 6,4e-84 Length: 1531
Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.26% Indels: 1
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-828-303-4 (1-1531)
QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGly 20
DB 655 ATGGCCACCGAGCGTGTCTTCAGAGACGACCTCGAGGCCCTCGAGGC 714
QY 21 CysIysAsnLeuCySGlyPhePheGlySerGlnAlaThrMetGlyLeuCy 40
DB 715 TGCAGAAACCTTTGGGGCTTCTTCGGCAGCCCAAGCTACCATGGGGTTGT 774
QY 41 TyrArgGluThrValMetGlnAlaIysMetThrAlaLeuAlaGluGlnAla 60
DB 775 TACCCAGAGACAGTCATGCACAGCG- AAGATGACGGCTTTAGCTGACAG 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProVal 80
DB 834 GCTCAGGCGACATCTGCCACAGCTGCTGCTTACGCCCCCGCTCTCTG 893
QY 81 LysLeuThrCySGluValGluArgThrMetIleValProHisGlnSerSer 100
DB 894 AAGCTCACATGCGAGGTTGAGAGAACATGATTTGCCGCATCATCTT 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValIysSer 120
DB 954 CAAGCCTGGTTACCCCGCTGACAGCTGCCCTCAGGACGTGAACTCT 1013
QY 121 ProSerArgProGluProAsnArgCySGlySerCySArgIysArgValG 140
DB 1014 CCTCTAGACCCGAGCCCAATGATCGGATCTTCAGAGAACGCTGTG 1073
QY 141 PheIysCySArgCySGlyAsnLeuTyrcySAlaLeuHisArgTyrcySer 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTACGCTTACATCGGTACTCGG 1133
QY 161 CysThrTyrcyAspTyrcyIysAlaAlaGlyGlnGluAlaIleAlaIysAla 180
DB 1134 TGCACATATGACTACAAAGCCGACAGGCGAGAAAGATTCGAAAGCTA 1193
QY 181 ValAlaGluIysValValIysPhe 188
DB 1194 GTGGCCGAGAGGTTGTCAAGTTT 1217

RESULT 3
US-09-244-805-10
; Sequence 10, Application US/09244805
; Patent No. 669660
; GENERAL INFORMATION:
; APPLICANT: Worley, Paul F.
; APPLICANT: Lanhuan, Anthony
; APPLICANT: Goetz, Bernard
; APPLICANT: Heilmisch, Holger
; APPLICANT: Kuner, Robini
; APPLICANT: Schiek, Sigrid
; APPLICANT: Nikolich, Karoly
; APPLICANT: Zhukovskii, Eugene
; TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 10496/004001
; CURRENT APPLICATION NUMBER: US/09/244,805
; CURRENT FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/074,518
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: 60/074,135
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 2515
; TYPE: DNA
; ORGANISM: Eukaryote
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (414)...(1055)

US-09-244-805-10	Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
	4.35e-26	Length:	2515				
	367.00	Matches:	81				
	53.27%	Conservative:	33				
	37.85%	Mismatches:	66				
	37.30%	Indels:	34				
	3	Gaps:	7				
US-10-716-089-20 (1-188) x US-09-244-805-10 (1-2515)							
QY	4	GIUARGVAlSerGIgluThrThrsSerGlnAlaProGIgluProValMetCysIysAsn	23				
DB	408	AAAAATATGGCTCAGAG--ACTAAACGAGACCCCA--GGGCCATGCTGTAGTACT	461				
QY	24	LeuCYsgIyPhePheGlySerGlnAlaThrMetGlyLeuCYsSerIyCyS/TYrArgGlu	43				
DB	462	GGATGTGGCTTTATGTGGAAATCCTAGACAAATGGAATGTCTGTCTTCTTCAAGAA	521				
QY	44	ThrValMetGlnAlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAla	63				
DB	522	CATCTTCAGAGACACAGAAATGTGGCAGAAATGAGCCCAATGGGAGCAGTAGGTTC	581				
QY	64	ThrsSerAlaThr-----AlaAlaAlaValGln-----	72				
DB	582	AACAGTCTCACTCAGACTCTGCTGTGTACAAAGAGCAGAGTCTACTTAAACACTGT	641				
QY	73	-----ProProAla-----Pro	76				
DB	642	GAGGTGCTGCTGGCAGCACATCTGAAAAATCAAGAAATGTGCTGTGGCTGCTTGCCT	701				
QY	77	ValHisGluThrIhrLeuLeuThrCysGluValGluArgThrMetIleValProHisGlnSer	96				
DB	702	GTAATCAACAATGACAGAAATGAGCATTTCAAGAGAGGACAAATTAACCTCCCGAAA	761				
QY	97	SerSerIyTyrGlnGlnAspLeuValThr-----ProAlaAlaAlaAlaProGlnAla	113				
DB	762	ACAGAGGTGTGACAGCCAGTGTCTCAGCCCACTCACTCACTCTCAGCCCACTTCT	821				
QY	114	ValIysSerSerIleAlaAlaProSerArgProGluPro-----AsnArgCysGlySer	131				
DB	822	TCTCAAAATGAGAAAGAAAGCTCTGAGTTGCCCAACCAAGAGAAACAGATGTTTATG	881				
QY	132	CysArgIyIysArgValGlyLeuThrGlyPheIysCysArgCysGlyIysAsnLeuTyrCysAla	151				
DB	882	TGTGAAAGAAAGTGGCTTACAGGCTTACAGGCTTACAGGCTTACAGGCTTACAGGCTT	941				
QY	152	LeuHisArgTyrSerAspIyHisIhrThrCysThrTyrAspTyrIysAlaAlaGlyGlnIu	171				
DB	942	CTTCACTCTTACTCTGACACAGCACTGCTCTTATGATTACAAACAGAGCTGACGA	1001				
QY	172	AlaIleAlaIysAlaAsnProLeuValValAlaGlyIysVal	185				
DB	1002	AAATCAGAAAGAAATTCAGTGTGTGCTGAAAAATTC	1043				
RESULT 4							
US-09-949-016-810							
Sequence 810, Application US/09949016							
Patent No. 6812339							
GENERAL INFORMATION:							
APPLICANT: VENER, J. Craig et al.							
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED							
FILE REFERENCE: CLO010107							
CURRENT APPLICATION NUMBER: US/09/949,016							
CURRENT FILING DATE: 2000-04-14							
PRIOR APPLICATION NUMBER: 60/241,755							
PRIOR FILING DATE: 2000-10-20							
PRIOR APPLICATION NUMBER: 60/237,768							
PRIOR FILING DATE: 2000-10-03							
PRIOR APPLICATION NUMBER: 60/231,498							
PRIOR FILING DATE: 2000-09-08							

	NUMBER OF SEQ ID NOS: 207012	
/	SOFTWARE: FastSeq For Windows Version 4.0	
/	SEQ ID NO 810	
/	LENGTH: 2425	
/	TYPE: DNA	
/	ORGANISM: Human	
/	US-09-949-016-810	
Alignment Scores:		
Pred. No.:	6,48e-26	Length: 2425
Score:	365.00	Matches: 81
Percent Similarity:	52.80%	Conservative: 32
Best Local Similarity:	37.85%	Mismatches: 67
Query Match:	37.09%	Indels: 34
Gaps:	3	7
US-10-716-089-20 (1-188) x US-09-949-016-810 (1-2425)		
QY	4 GlnArgValSerGlnGluThrThrSerGlnAlaProGlnGlyProValM ::: ::::: :::::	23
Db	283 AAAAATATGGCTCAGGAG---ACTAACCAACAGCCCG--GGGCCCATGC ::: ::::: :::::	336
QY	24 LeuCysglYPhePheGlySerGlnAlaThMetGlyLeucYSerSylSC ::::: :::::	43
Db	337 GGATGGGCTTTATTGGAAATCTCAGACCAATGAATGAATGGTGTCACTTG ::::: :::::	396
QY	44 ThrValMetGlnAlaIaySwetHraIalaeuAlaglunAlaThrgInA ::: ::: :::::	63
Db	397 CACTTCAGAGCGCACCAATAATAGCGAGATGAGACCCTATGGGACAG ::: ::: :::::	456
QY	64 ThSerLathr-----AlaalaaIaValGlN----- ::::: :::::	72
Db	457 AACAGTCTACTCTCAGATTCGATCTGTACAGAGACAGACACTAGCT ::::: :::::	516
QY	73 -----ProProA -----	76
Db	517 GAAGGTCCTCGGACGACATCTGAAAAATCAAAGAATGCTCTGTGG -----	576
QY	77 ValHisgluThrLybUthrCySGluValaGluaGrThrMetileValP ::: ::: ::::: :::::	96
Db	577 GTNACTCAGCAANTGACAGAAATGAGCATTTCAAAGAGGACAAAATTA ::: ::: ::::: :::::	636
QY	97 SerSerTyGlnGlnAspleuValThr-----ProAlaAlaIaA ::: ::: ::::: :::::	113
Db	637 ACAGAGGTGTCAGAGCCAGTGTGCACCTCAGCCCAGTCAATCAGTTCTC ::: ::: ::::: :::::	696
QY	114 ValLySerSerlelalaIaProSerArgProGluPro-----AsnA ::: ::: ::::: :::::	131
Db	697 TCTCAGAGTGMAAGAAAAAGCTCTGAATTGCCCAACCAAGAAAAACA ::: ::: ::::: :::::	756
QY	132 CyArGlybArGValaGlyLeuthrGlyPhelYCyArGCSyGlybent ::::: :::::	151
Db	757 TGCGAAGAAAGAACTTGCTCTTACAGGTTTGAACTGCCGATGTGAAAT ::::: :::::	816
QY	152 LeuHisarGTYSerAsplyShisThrCYeThrTYrAspyrLYsaIaA ::::: :::::	171
Db	817 CTTCACCGTTACTGTGACAAAGCACAACTGCCGTATGATTACAAAGAC ::::: :::::	876
QY	172 AlaIlleAlaIyaIaIsenProdeuValaIlaGlulYsVal 185 ::::: :::::	
Db	877 AAATCAGAAAAAGAAATCCAGTTGTGTGGCTGAAAAAATT 918 ::::: :::::	
RESULT 5		
US-09-949-016-3693		
Sequence 3693, Application US/09949016		
Patent No. 6812339		
GENERAL INFORMATION:		
APPLICANT: VENTER, J. Craig et al.		
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED		
WITH TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTIC		
FILE REFERENCE: CL001307		
CURRENT APPLICATION NUMBER: US/09/949, 016		
CURRENT FILING DATE: 2000-04-14		
THREE		

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3693
; LENGTH: 2684
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3693

Alignment Scores:
Pred. No.: 7,44e-26 Length: 2684
Score: 365.00 Matches: 81
Percent Similarity: 52.80% Conservative: 32
Best Local Similarity: 37.85% Mismatches: 67
Query Match: 37.09% Indels: 34
DB: 3 Gaps: 7

US-10-716-089-20 (1-188) x US-09-949-016-3693 (1-2684)
QY 4 GluATGValSerGlnGluThrThrSerGlnAlaProGluGlyProValMetCysIysAsn 23
DB 542 AAAAATATATGGCTCAGAGAG---ACTAACCAAGACCCG---GGGCCATGCTGTGTAGCACA 595
QY 24 LeuGysGlyPheGlySerGlnAlaThrMetGlyLeuGysSerIysCysTyrArgGlu 43
DB 596 GGATGTGGCTTTATGAAATCTCTAGCAAAATGAAATGTGTTGATTGCTACAAAGAA 655
QY 44 ThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAla 63
DB 656 CATCTTCAGAGGCAGCAAAATATGAGCAGAAATGAGCCCAATGGGAGCAGCTAGTGGTTCC 715
QY 64 ThrSerAlaThr-----AlaAlaAlaValGln----- 72
DB 716 AACACTCTACTCTCAGATTCTGCTATCTGTACAGAGCAGACACTTAACAACACTGT 775
QY 73 -----ProProAla-----Pro 76
DB 776 GAAGGTGCTGCTGAGCAGACATCTGAAATATCAGAAATATGTCCTGTGGCTGCTTGGCT 835
QY 77 ValHisGluThrLysLeuThrCysGluValAlaGluThrMetIleValProHisGlnSer 96
DB 836 GTAATCTCAGCAAAATATGACAAATGAGCATTTCAGAGAGGACAAATATTAATACCCGAAA 895
QY 97 SerSerTyrGlnGlnAspLeuValThr-----ProAlaAlaAlaAlaProGlnAla 113
DB 896 ACAGAGGTGTCTCAGACCAAGTTCTCACTCAGCCCAATCCATCAGTTTCTCAGCCAGTACT 955
QY 114 ValLysSerSerIleAlaAlaProSerArgProGluPro-----AsnArgCysGlySer 131
DB 956 TCTCAGAGTGAAGAAAAGCTCTCTGAATTCGCCAAACCAAGAAAACAGAGTTTCATG 1015
QY 132 CysArgLysArgValGlyLeuThrGlyPheLysCysArgCysGlyAsnLeuThrCysAla 151
DB 1016 TGCAGAAAAGAAATGTGTCTTACAGGGTTTGACTGCCAGTGTGGAAATTTGTTGTGGA 1075
QY 152 LeuHisArgTyrSerAspLysHisThrCysThrTyrAspTyrLysAlaAlaGluGlnGlu 171
DB 1076 CTTTCCCGCTTACTCTGACAGACCAACTGTCGTTATGATTACAAAGCAGAAAGCTGCAGCA 1135
QY 172 AlaIleAlaLysAlaAsnProLeuValValAlaGluLysVal 185
DB 1136 AAAATCAGAAAAGAGATCCAGTGTGTGTGGCTGAAAAAATT 1177

RESULT 6
US-08-861-269-4
; Sequence 4, Application US/08861269
; Patent No. 5817494
; GENERAL INFORMATION:
```

```

; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,269
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1197 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT09
; CLONE: 2151473
; US-08-861-269-4

Alignment Scores:
Pred. No.: 2.48e-24 Length: 1197
Score: 344.50 Matches: 81
Percent Similarity: 51.14% Conservative: 31
Best Local Similarity: 36.99% Mismatches: 67
Query Match: 35.01% Indels: 40
DB: 2 Gaps: 7

US-10-716-089-20 (1-188) x US-08-861-269-4 (1-1197)
QY 2 AlaThrGlu-ArgValSerGlnGluThrThrSerGlnAlaProGluGlyI 21
DB 249 GCAACTGAGGAACATGGCTCAAGAAACTAATACACAGC-----CAAGTGC 302
QY 21 slyAsnLeuGlyCysGlyPheGlySerGlnAlaThrMetGlyLeuGysI 41
DB 303 TTCCACTGGCTGTGGATTATGGAACCTCGTACAAATGAGCAGTGT 362
QY 41 rArgGlu----- 45
DB 363 TAAAGAACATCTTCAAGACAGAAATGATGTAATGTAGTAATAGCCCA 422
QY 45 lMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThr----- 60
DB 423 TGTCAGTAGTCTGCTGATCTTACCACTTCAATGACAGATGGCAGT 482
QY 60 aAlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaPro 80
DB 483 CCACTCAGCATTTAAGACTTACATCTTCAATCTTAATGAG---CCAGCCCT 539
QY 80 rLysLeuThrCysGluValGluArgThrMetIleValProHisGlnSer 100
```



```

Db      540 GTCACCTTTATTCGAA-----TCGTAGCATCTTCAATTGGACGTCATC 587
Qy      100 nglnapleuvalthrproalaalaalaproglinala---vallyserse1lea1 119
      588 TGTGACAAAGCAGTACCTGAAACAGAAAGATGTGCAGGGCTTCACTTACAGACACAC 647
Qy      119 aalapraser-----ArgProgluProabnrrcy 129
Db      648 GCAGCCATCTGAAGACAAAGCAAGTCTCTTGAACCGAACAACAAAGATGCTG 707
Qy      129 egllyserCyarglyarGValglyleuThrglyPheLyCyargCyeglYasnleuty 149
Db      708 TTTCACTGTCAGAGAAAGAAAGTGGACTTACTGSGTTTGAATCCSGTGTGAAATGTTTA 767
Qy      149 rCyalaLeuHlsarGlyrSerapLyshIsthrcYsthryrAspyrlyeAla1a1a1 169
      768 CTGTGCTGTACACCGTTACTCAGATGTACACATTTCTTACATTACAAAGCCGATGC 827
Qy      169 yglnglnuallealalyalaasnProleuValVala1a1a1a1a1a1a1a1a1a1 187
Db      828 TGTGTGAAATTCAGAAAGAAATCCAGTAgTGTGTGAAAGATCCAAAG 882

RESULT 7
US-09-134-596-4
; Sequence 4, Application US/09134596
; Patent No. 5922318
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,596
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/861,269
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1197 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT09
; CLONE: 2151473
; US-09-134-596-4

Alignment Scores:
Pred. No.: 2,48e-24 Length: 1197
Score: 344.50 Matches: 81

```

```

Percent Similarity: 51.14%
Best Local Similarity: 36.99%
Query Match: 35.01%
DB: 2
Gaps: 7

US-10-716-089-20 (1-188) x US-09-134-596-4 (1-1197)

Qy      2 Alathrglu-ArgvalSerGlnGlnuThrThrserrlnalaprogluGly 21
      249 GCAACTGAGGAAACATGCTCAAGAAACTAATCAACAGC-----CAAGTC 302
Qy      21 elYasnleuCyarglyPhePheGlySerGlnalathrmetGlyleuCy 41
Db      303 TTTCACTGCTGTGGATTTATGAAACCTCGTCAAAATGGCAGTGT 362
Qy      41 rArglu----- 45
Db      363 TAAAGACATCTTCAAGACAGATAGTAGTAATGTAATAGGCCCA 422
Qy      45 lmetGlnalalysermetThrAlaLeuAlaGlnAla1a1a1a1a1a1a1a1a1a1 60
Db      423 TGTCACTAGTCTGTGTGATCTTTACAGTTCAATGCACAGATGCAGT 482
Qy      483 CCAAGTCAGCATTAGACTCTTCACTTCACTATAGCAG---CCAGCCCT 539
Db      80 rlyseuThrcYsglyuValgluArgThrmetileValProhIgsnser 100
      540 GTCACCTTTATTCGAA-----TCGTAGCATCTTCAATTG 587
Qy      100 nglnapleuvalthrproalaalaalaproglinala---vallys 119
      588 TGTGACAAAGCAGTACCTGAAACAGAAAGATGTGCAGGCTTCAGTATCA 647
Qy      119 aalapraser-----ArgProglu 129
Db      648 GCAGCCATCTGAAGACAAAGCAAGTCTCTTGAACCGAACAACAAAG 707
Qy      129 egllyserCyarglyarGValglyleuThrglyPheLyCyargCy 149
Db      708 TTTCACTGTCAGAGAAAGAAAGTGGACTTACTGSGTTTGAATCCSGTGT 767
Qy      149 rCyalaLeuHlsarGlyrSerapLyshIsthrcYsthryrAspyrly 169
      768 CTGTGCTGTACACCGTTACTCAGATGTACACATTTCTTACATTAC 827
Qy      169 yglnglnuallealalyalaasnProleuValVala1a1a1a1a1a1a1a1a1a1 187
Db      828 TGTGTGAAATTCAGAAAGAAATCCAGTAgTGTGTGAAAGATTC

RESULT 8
US-09-293-273-4
; Sequence 4, Application US/09293273
; Patent No. 6057112
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS

```

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/293,273
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,596
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-293-273-4

Alignment Scores:

Pred. No.:	2,486-24	Length:	1197
Score:	344.50	Matches:	81
Percent Similarity:	51.14%	Conservative:	31
Best Local Similarity:	36.99%	Mismatches:	67
Query Match:	35.01%	Indels:	40
DB:	3	Gaps:	7

US-10-716-089-20 (1-188) x US-09-293-273-4 (1-1197)

```
QY 2 Alathrglu-ArgvalserglnluThrThrserglnalaprogluglyprovalmetCy 21
DB 249 GCAACTGAGGAACATGGCTCAAGAACTAATCAGAC-----CAAGGCGCTATGCTTGG 302
QY 21 glyvalnleucyrglypneheglyserglnalathrmetglyleucyserlyserCy 41
DB 303 TTCACCTGGCTGTGGATTTTATGAAACCTCGTCAATGGCATGTGTTCAGTATGCTA 362
QY 41 rarglu-----ThrVa 45
DB 363 TTAAGAACCTCTTCAAGACAGATATGTAATGTAGTAATACCCACCTGCACCTC 422
QY 45 lmerglnalalywecthrAlaLeuAlaGluGlnAlathr-----GlnAl 60
DB 423 TGTCAGTAGTGTCTGATCTTTTACAGTTCAATGCACAGATGGCAGTGTCCAGAAAGC 482
QY 60 aalaelnathrserAlathrAlaAlaAlaValGlnProProAlaprovalhieglnuth 80
DB 483 CCAGTCACACTTAGACTTACATCTTCACTCATGTGAG-----CCACCCCTGTATCAATCA 539
QY 80 rlyseruthrCyrgluValGluArgThrmetileValProhIsGlnserSerSertygl 100
DB 540 GTCACTTTTATAGAA-----TCTGTAGCATCTTCTCAATTGGACAGTACATC 587
QY 100 nglnaspleuValThrProAlaAlaAlaAlaProGlnAla---VallyserSerileAl 119
DB 588 TGTGACAAAGACAGTACCTGAAGAAAGATGTGACAGGCTTCAGATATCAGACACAGACA 647
QY 119 aAlaProser-----ArgProGluProhArgCy 129
DB 648 GCAGCCACTCTGAAGACAAAGCAAGTCTCTTGAACAAACCGAAACAAACAAATCGCTG 707
QY 129 gglYserCyrglyValGlyLeuThrglyPheLysCyrglyValAsnLeuTy 149
DB 708 TTTCATGTGCGAGGAAGAAAGTGGACTTACTGGGTTTGAATCCCGGTGGAAATGTTTA 767
QY 149 rCyvalaleuHlsarGlySerAspLysHlsThrCyserThryraapTyrylysalAlaG 169
DB 169
```

```
DB 768 CTGTGTGTACACCGCTTACTCAGATGTACACATTCGCTTACATATAC
QY 169 yglngluAlaIleAlaLysAlaAsnProLeuValAlaGlnLysVal
DB 828 TGTGAGAAATCAGAAAGAAATCCAGTATGTTGTGTGTAAGATCC
```

RESULT 9

US-09-640-211A-1804
Sequence 1804, Application US/09640211A
Patent No. 6833446
GENERAL INFORMATION:
APPLICANT: Wood, Marion
APPLICANT: Shenk, Michael A.
APPLICANT: McGrath, Annette
APPLICANT: Glenn, Matthew
TITLE OF INVENTION: Compositions and Methods for the
FILE OF INVENTION: Modification of Gene Transcription
FILE REFERENCE: 11000.1021CIU
CURRENT APPLICATION NUMBER: US/09/640,211A
CURRENT FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 2368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1804
LENGTH: 533
TYPE: DNA
ORGANISM: Pinus radiata
US-09-640-211A-1804

Alignment Scores:

Pred. No.:	3,16e-18	Length:	533
Score:	277.00 <td>Matches:</td> <td>19</td>	Matches:	19
Percent Similarity:	54.25% <td>Conservative:</td> <td>64</td>	Conservative:	64
Best Local Similarity:	41.83% <td>Mismatches:</td> <td>46</td>	Mismatches:	46
Query Match:	28.15% <td>Indels:</td> <td>24</td>	Indels:	24
DB:	3 <td>Gaps:</td> <td>5</td>	Gaps:	5

US-10-716-089-20 (1-188) x US-09-640-211A-1804 (1-533)

```
QY 5 Argvalserglnlu-----ThrThrsrg 16
DB 114 AGGATGATCAAGAAATCTGAACATCGAGCTGATGCGACTGCTCC 173
QY 17 glyprovalmetCylyAsnLeuCyrglypneheglyserglnalathr 36
DB 174 GGGCACTCTTTCGGCAATACATCGCGCTTTTGGCAGTGGCAAC 233
QY 37 CyserlyserCyryrargluThrValmetGlnAlalywecthrAla 56
DB 234 TGTTCGAAATGTTACAGGATCTGATTAAGA-----GAGGCCCT 284
QY 57 AlathrglnAlaGlnAlaThrSerAlathrAlaAlaValGlnP 76
DB 285 GCAATGGCCGCGTTGAGAGTCAATTGCGCGGCTTCCGATGGAG 344
QY 77 ValhieglnuthrLysLeuThrCyrgluValGluArgThrmetileValP 93
DB 345 CTTTCCAAAGCAGATGTTTGTGTGAACAAGCCGTGACCGATCTCC 404
QY 94 HlsGlnserSerSertyrglnGlnaspleuValThrProAlaAlaAla 113
DB 405 CAAGCTGTGACGTTCACTTGGCTGATATA-----GGTTATCTTCTT 455
QY 114 VallyserSerileAlaAlaProserArgProGluProhArgCy 133
DB 456 -----CTTCTGCGAAACTCTTAACCGGTGCT 494
QY 134 LysArgValGlyLeuThrglyPheLysCyrglyValAsnLeuTy 146
DB 495 AAACGATCGGTCTGACCGCTTCAATATGTGCGTGGCA 533
RESULT 10
US-09-313-294A-7446
Sequence 7446, Application US/09313294A
```

```
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Laljudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 7446
/ LENGTH: 283
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700381739H1
/ LOCATION: 216
/ NAME/KEY: unsure
/ OTHER INFORMATION: a, t, c, g, or other
/ US-09-313-294A-7446

Alignment Scores:
Pred. No.: 1,986-17 Length: 283
Score: 265.00 Matches: 44
Percent Similarity: 85.48% Conservative: 9
Best Local Similarity: 70.97% Mismatches: 9
Query Match: 26.93% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-313-294A-7446 (1-283)

QY 126 ProaenargCyeglySerCySaRgLySaRgValGlyLeuThGlyPheLyScySaRgCys 145
DB 2 CCAACCGGTCGCGCAACCTGTAGGAAGCGTGTGGGTTTAACTGCCGATGC 61
QY 146 GlyaenLeuTYrCySaAlaLeuHISaRgTYrSerAaPlySHISrThrCyST 165
DB 62 GGGAAcAcTgTAcTgTTCGATGcAcCGcTAcTCCGAAcAcAcGAcTGCAGTTCGACTAT 121
QY 166 LysAlaAlaGlyGlnGlnAlaIleAlaIleAlaLysAlaAsnProLeuValValAlaGlnLysVal 185
DB 122 CCACTGCAGCTAGGAGCAGCTATCCGCAAGGCCAATCCAGTGTGAAGCGAGAAAGCTT 181
QY 186 ValLys 187
DB 182 GACAAG 187

RESULT 11
US-09-313-294A-1951
/ Sequence 1951, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Laljudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 1951
/ LENGTH: 281
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700551796H1
/ US-09-313-294A-1951

Alignment Scores:
```

```
Pred. No.: 4,816-17 Length: 281
Score: 261.00 Matches: 43
Percent Similarity: 87.10% Conservative: 11
Best Local Similarity: 69.35% Mismatches: 8
Query Match: 26.52% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-313-294A-1951 (1-281)

QY 126 ProaenargCyeglySerCySaRgLySaRgValGlyLeuThGlyPheLy 145
DB 10 CCAAGCCGGTCGCGCGCCCTGCAGGAAGCGTGTGGGCTGACGGATTTT 69
QY 146 GlyaenLeuTYrCySaAlaLeuHISaRgTYrSerAaPlySHISrThrCyST 165
DB 70 GGGAAcAcTgTAcTgTTCGATGcAcCGcTAcTCCGAAcAcAcGAcTGCAGTTCGACTATC 129
QY 166 LysAlaAlaGlyGlnGlnAlaIleAlaIleAlaLysAlaAsnProLeuValVal 185
DB 130 CCGACTGCAGCAAGGAGCAGCTATGCCAAGGCCAATCCGTGTGAGG 189
QY 186 ValLys 187
DB 190 GACAAG 195

RESULT 12
US-09-313-294A-3856
/ Sequence 3856, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Laljudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVE
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 3856
/ LENGTH: 211
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700282265H2
/ NAME/KEY: unsure
/ LOCATION: 41, 60, 99, 116, 118, 135, 154
/ OTHER INFORMATION: a, t, c, g, or other
/ US-09-313-294A-3856

Alignment Scores:
Pred. No.: 4,066-13 Length: 211
Score: 219.00 Matches: 39
Percent Similarity: 79.37% Conservative: 11
Best Local Similarity: 61.90% Mismatches: 12
Query Match: 22.26% Indels: 1
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-313-294A-3856 (1-211)

QY 126 ProaenargCyeglySerCySaRgLySaRgValGlyLeuThGlyPheLy 145
DB 9 CCAAGCCGGTCGCGCGCCCTGCAGGAAGCGTGTGGGCTGACGGGATTTA 68
QY 146 GlyaenLeuTYrCySaAlaLeuHISaRgTYrSerAaPlySHISrThrCyST 165
DB 69 GGGAAcAcTgTAcTgTTCGATGcAcCGcTAcTCCGAAcAcAcGAcTGCAGTTCGACTATC 128
QY 165 LysAlaAlaGlyGlnGlnAlaIleAlaIleAlaLysAlaAsnProLeuValVal 185
DB 129 TCGACACGCAAGGAGCAGCGGATTCGCAAGGCCAATCCGTGTGAGG 188
```

APPLICANT:	Xue, Aiyong J.	
APPLICANT:	Yang, Yonghong	
APPLICANT:	Wang, Jian-Rui	
APPLICANT:	Zhou, Ping	
APPLICANT:	Ma, Yungting	
APPLICANT:	Wang, Dunrui	
APPLICANT:	Wang, Zhiwei	
APPLICANT:	John Tillinghast	
APPLICANT:	Drmanac, Radoje T.	
TITLE OF INVENTION:	No. 656962e21 Nucleic Acids and	
TITLE OF INVENTION:	Polypeptides	
FILE REFERENCE:	784CIP2B	
CURRENT APPLICATION NUMBER:	US/09/620,312D	
PRIOR FILING DATE:	2000-07-19	
PRIOR FILING DATE:	2000-04-25	
PRIOR APPLICATION NUMBER:	09/488,725	
PRIOR FILING DATE:	2000-01-21	
NUMBER OF SEQ ID NOS:	1105	
SOFTWARE:	PC_Fl_genes Version 1.0	
SEQ ID NO 849		
LENGTH:	1281	
TYPE:	DNA	
ORGANISM:	Homo sapiens	
FEATURE:		
NAME/KEY:	CDS	
LOCATION:	(416)..(1033)	
US-09-620-312D-849		
Alignment Scores:		
Pred. No.:	1,45e-09	Length: 1281
Score:	193.50	Matches: 50
Percent Similarity:	43.09%	Conservative: 31
Best Local Similarity:	26.60%	Mismatches: 72
Query Match:	19.66%	Indels: 35
DB:	3	Gaps: 7
US-10-716-089-20 (1-188) x US-09-620-312D-849 (1-1281)		
QY	12 SerGlnAlaProGluGlyProValMetCysIlyuAsnLeuCyseGlyPheP	31
DB	440 AGCAAAAGCGCCAGCGCTGCGGCTGCGTCC-----TGGGGCTTC	493
QY	32 AlaThrMetGlyLeuCyseSerIlyeCyseIyIyArgGlu-----	43
DB	494 AAGACTATGAATCTCTGTTCCAAATGCTTTGCTGATTTTCAAAGAAC	553
QY	44 -----ThrValMetGlnAlaIlyseThrAlaLeuAlaGluG	59
DB	554 GATTCGCGCTCAAGTACAAAGTAACAGCCCAATCAGATTGTTTCCGAG	613
QY	60 AlaAlaGlnAlaThrSerAlaThrAlaAlaAlaValAlaGlnPro-----	76
DB	614 GACAAACAACATACCTCGATACACACGCCCACTCTTAGTCCAGCAGC	673
QY	77 ValHisGluThrIlyeLeuThrCyseGluValGluIyArgThrMetIlyeAlP	96
DB	674 -----ACAGAAACGAAATGTAATTCAC	706
QY	97 SerSerIyGlnGlnAspLeuValThrProAlaAlaAlaAlaProGlnA	116
DB	707 GAGTATTCACAGCTGAGAAATGAGGCTTCACCAAGTAACAGCCAGACT	766
QY	117 Ser-----IleAlaAlaProSerArgProGluProAsnA	131
DB	767 ACGGAACGCTCCGAGGAACCAAGTCAGATCTAAACAGAGAGTCGACGTC	826
QY	132 CysArgIlyArgValGlyIlyeThrGlyPheIlye-----CysArgCy	148
DB	827 TGGCAAAACCAATCGAGACTGGTCAGACAGGAATGGAGTCGTGCTC	886
QY	149 TyrCyAlaLeuHisArgIySerAspIlyHisIthrCysIthTyrAspIy	168

Db 887 TCTGTATGTATCATCGCCTCCCGAGCAGCAGTGCATTCGACCATGSGCGT 946
Qy 169 GY---GNGLUAlleAllys 175
Db 947 GCGCGGAGAGCCATCATGAAA 970

RESULT 15
US-09-949-016-12552
/ Sequence 12552, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12552
/ LENGTH: 14524
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-12552

Alignment Scores:
Pred. No.: 4.39e-08 Length: 14524
Score: 193.00 Matches: 33
Percent Similarity: 76.60% Conservative: 3
Best Local Similarity: 70.21% Mismatches: 11
Query Match: 19.61% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-949-016-12552 (1-14524)

Qy 139 ThrGlyPheLeCyArgCyAGlyAsnLeuTyrcysAlaLeuHlaArgTySerAspLys 158
Db 10877 ACAGGTTTGACTGCCGAGTGGAAATTTGTTGTGACTTCACCGTTACTCTGACAG 10936
Qy 159 HleThrCyethrTyAspTyrlLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnPro 178
Db 10937 CACAACGTGCGTATGATTACAAAGCAGAACTGCAGCAAAATCAGAAAGAGATCCA 10996
Qy 179 LeuValIvalAlaGluLysVal 185
Db 10997 GTTGTGTGGCTGAAAAAATT 11017

Search completed: December 8, 2005, 18:44:31
Job time : 172 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:36:37 ; Search time 791 Seconds
(without alignments)
1965.415 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATERVSQRTTSGABGPMV.....GGEAIKANKPLVAKKVKVF 188

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cg2_1/USPTO_spool_p/US10716089/crunat_05122005_094344_13199/app_query.fasta_1.327
-DB=Published Applications NA.Main -QFMT=faetap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=spct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10716089 @CGN 1 1549 @runat_05122005_094344_13199 -NCPU=6 -ICPU=3
-NO WMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.Main:*

- 1: /cg2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cg2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cg2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cg2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 5: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 6: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 7: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 8: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 9: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 10: /cg2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	984	100.0	804	3	US-09-828-303-12
2	984	100.0	804	7	US-10-716-089-12
3	957	97.3	1531	3	US-09-828-303-4
4	957	97.3	1531	7	US-10-716-089-4
5	425.5	43.2	945	7	US-10-425-114-26146
6	425.5	43.2	947	8	US-10-425-115-71431
7	425.5	43.2	1142	7	US-10-767-701-15574

8	419.5	42.6	2671	7	US-10-437-963-21113	Seq	/	A
9	416.5	42.3	1032	7	US-10-425-114-27574	Seq	/	A
10	416.5	42.3	1171	7	US-10-437-963-37809	Seq	/	A
11	416.5	42.3	1249	7	US-10-437-963-37808	Seq	/	A
12	415	42.2	736	7	US-10-425-114-22047	Seq	/	A
13	415	42.2	776	7	US-10-425-114-28524	Seq	/	A
14	415	42.2	821	7	US-10-425-114-28524	Seq	/	A
15	415	42.2	826	7	US-10-425-114-21992	Seq	/	A
16	415	42.2	841	7	US-10-425-114-25691	Seq	/	A
17	415	42.2	846	7	US-10-425-114-25727	Seq	/	A
18	415	42.2	853	7	US-10-425-114-28785	Seq	/	A
19	415	42.2	863	7	US-10-425-114-27172	Seq	/	A
20	415	42.2	863	7	US-10-425-114-28540	Seq	/	A
21	415	42.2	868	7	US-10-425-114-15838	Seq	/	A
22	415	42.2	868	7	US-10-425-114-28768	Seq	/	A
23	415	42.2	872	7	US-10-425-114-26442	Seq	/	A
24	415	42.2	872	7	US-10-425-114-27856	Seq	/	A
25	415	42.2	873	7	US-10-425-114-27861	Seq	/	A
26	415	42.2	876	7	US-10-425-114-27629	Seq	/	A
27	415	42.2	886	7	US-10-425-114-6594	Seq	/	A
28	415	42.2	894	7	US-10-425-114-23540	Seq	/	A
29	415	42.2	898	7	US-10-425-114-32027	Seq	/	A
30	415	42.2	899	7	US-10-425-114-25461	Seq	/	A
31	415	42.2	902	7	US-10-425-114-27516	Seq	/	A
32	415	42.2	904	7	US-10-425-114-21878	Seq	/	A
33	415	42.2	918	7	US-10-425-114-32045	Seq	/	A
34	415	42.2	920	7	US-10-425-114-22547	Seq	/	A
35	415	42.2	924	7	US-10-425-114-31655	Seq	/	A
36	415	42.2	926	7	US-10-425-114-27617	Seq	/	A
37	415	42.2	928	7	US-10-425-114-27609	Seq	/	A
38	415	42.2	931	7	US-10-425-114-28237	Seq	/	A
39	415	42.2	932	7	US-10-425-114-25774	Seq	/	A
40	415	42.2	942	7	US-10-425-114-28518	Seq	/	A
41	415	42.2	943	7	US-10-425-114-28353	Seq	/	A
42	415	42.2	948	7	US-10-425-114-5385	Seq	/	A
43	415	42.2	953	7	US-10-425-114-27813	Seq	/	A
44	415	42.2	956	7	US-10-425-114-27789	Seq	/	A
45	415	42.2	959	7	US-10-425-114-22492	Seq	/	A

ALIGNMENTS

RESULT 1
US-09-828-303-12
Sequence 12, Application US/09828303
Patent No. US20020102695A1
GENERAL INFORMATION:
APPLICANT: COSTA E SILVA, OSWALDO DA
APPLICANT: BOHNETT, HANS J.
APPLICANT: VAN THIELEN, NOCHA
TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROT
FILE REFERENCE: 16313-0030
CURRENT APPLICATION NUMBER: US/09/828,303
PRIOR FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/196,001
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 804
TYPE: DNA
ORGANISM: Physcomitrella patens
US-09-828-303-12

Alignment Scores:
Pred. No.: 4.78e-105
Score: 984.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 3
Length: 804
Matches: 188
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-716-089-20 (1-188) X US-09-828-303-12 (1-804)

QY	1	MetAla2ThrGlnAbxgrAlaSerGlnGluIuThrPheSerGlnAlaProGluGluProValMet	20
Db	134	ATGAGCCACCGAGGGTGTGTCTCAAGAGACAGACTTGCAAGGCCCCGTGAAGGGTCCAGTTATG	193
QY	21	CysAlaPheAsnLeuCyseGluPhePheGlySerGlnAlaThrMetGlyLeuCyseSerLeuCys	40
Db	194	TGCAGAAGACTTTGGGGGCTTCTTTGGCAGCCAGGTACATGGGGTTGTGTCCAAATGC	253
QY	41	TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla	60
Db	254	TACCGAGAGACACTGTATGCAGAGCAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCT	313
QY	61	AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHisGluThr	80
Db	314	GCTCAGGCGCAATCTTCACAGCTGTGTGTTCAGCCCCCGCTCCGTACATGAAGACC	373
QY	81	LysLeuThrCysGluValGluArgThrMetCysLeuAlaProHisGlnSerSerSerTyrGln	100
Db	374	AACCTCACATGGCAGGTTGAGAGAACAAATGATTTGGCCGATCAATCTTCCAGCTATCAA	433
QY	101	GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerLeuAlaAla	120
Db	434	CAAGACCTGGTTACCCCCCTGCAGCTCCCTCAGGACAGTGAAGTCTCTATGCGAGCT	493
QY	121	ProSerArgProGluProAsnArgCysGlySerCysValGlyValGluValGlyLeuThrGly	140
Db	494	CCCTCTAGACCCGAGGCCAATGCATGGCGATTTTCAGAGAAAGCGTTGGATTACAGGA	553
QY	141	PheLysCysArgCysGluLysAsnLeuTyrCysAlaLeuHisArgTyrSerAspLysHisThr	160
Db	554	TTTAAAGTGTGCTGTGGCAACTCTACTGCGTTTACATCGGTACTCGGAAACACACT	613
QY	161	CysThrTyrAspTyrLysAlaAlaGluGlnGluAlaLeuAlaLysValAsnProLeuVal	180
Db	614	TGCACATATATGACATAAAAGCCGAGGGCAGGAAGCGATTGGAAAGACTTAATCTCTTGTC	673
QY	181	ValAlaGluLysValValLysPhe	188
Db	674	GTGGCCGAGAAAGGTGTCAAGTTT	697

RESULT 2

US-10-716-089-12
: Sequence 12, Application US/10716089
: Publication No. US20040107463A1
: GENERAL INFORMATION:
: APPLICANT: COSTA E SILVA, OSWALDO DA
: APPLICANT: BOHNETT, HANS J
: APPLICANT: VAN THIELEN, NOCHA
: APPLICANT: CHEN, RUTYING
: TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
: FILE OF INVENTION: METHODS OF USE IN PLANTS
: FILE REFERENCE: 16313-0030
: CURRENT APPLICATION NUMBER: US/10/716,089
: CURRENT FILING DATE: 2003-11-18
: PRIOR APPLICATION NUMBER: 60/196,001
: PRIOR FILING DATE: 2000-04-07
: NUMBER OF SEQ ID NOS: 79
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 12
: LENGTH: 804
: TYPE: DNA
: ORGANISM: Physcomitrella patens
US-10-716-089-12

Alignment Scores:		
Pred. No.:	4.78e-105	Length: 8040
Score:	984.00	Matches: 188
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0

DB:	7	Gaps:	0
-----	---	-------	---

US-10-716-089-20 (1-188) X US-10-716-089-12 (1-804)

	Protein	Accession	Length
OY	Meth1athrGluArgValSerGlnGluThrThrsSerGlnAlaProGluGly	1	20
Db	ATGGCCACCGAGCGTGCTGTCTCAGAGAAGCAAGCTTCGAGGCCCTTGAAG	134	193
OY	CysIysAsnLeuCySGlyPhePheGlySerGlnAlaThrmcGlyLeuC	21	40
Db	TGCAGAAACCTTTGCGGCTTCTTGCGAGCCAACTACATGGGGTTGT	194	253
OY	TyrArgLutThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnA	41	60
Db	TACCGAGAGACATCATGCAAGCGAAMATGACGGCTTTAGCTGAGCAAG	254	313
OY	AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProV	61	80
Db	GCTAGGCGACATCTGCCACAGCTGCTGTGTTACGCCCCCGCTCTCG	314	373
OY	LysLeuthrCysGluValAlaGluArgThrmcIleValProHisGlnSers	81	100
Db	AAGCTCACATGCCAGGTGTGAGAGAACATAATTGTGCCGATCATCTTT	374	433
OY	GlnAspleuValThrProAlaAlaAlaAlaProGlnAlaValLysSers	101	120
Db	CMAACCTGGTTACCCCGCTGCAGCTGCCCTCAGACAGAAAGTCTCT	434	493
OY	ProserArgProGluProAsnArgCysGlySerCysArgLysArgValG	121	140
Db	CCCCTTAGACC CGAGCCCAATCGATGCGAATCTTGCAAGAAAGCGTGTTC	494	553
OY	PheLysCyBarGySGlyAsnLeuTyrcysAlaLeuHisArgTYrSera	141	160
Db	TTTTAAGTGTGCTGTGGCAACTTACTGCGCTTTACATCGTACTCGG	554	613
OY	CysThrTyraepTYrLysValalaGlyGlnGlnAlaIlealalySalaA	161	180
Db	TGCACATATGACTACAAGACCGCAGGCGAAGAAACGATTGCGAAAGCTA	614	673
OY	ValAlaGluLysValValValLysPhe	181	
Db	GTGGCCGAGAAAGTTGTCAAAGTTT	674	

RESULT 3

US-09-828-303-4
 ; Sequence 4, Application US/09828303
 ; Patent No. US20020102695A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COSTA E SILVA, OSWALDO DA
 ; APPLICANT: BOHNER, HANS J.
 ; APPLICANT: VAN THIELEN, NOCHA
 ; APPLICANT: CHEN, ROUYUNG
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEIN
 ; TITLE OF INVENTION: METHODS OF USE IN PLANTS
 ; FILE REFERENCE: 16313-0030
 ; CURRENT APPLICATION NUMBER: US/09/828,303
 ; CURRENT FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/196,001
 ; PRIOR FILING DATE: 2000-04-07
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 1531
 ; TYPE: DNA
 ; ORGANISM: Physcomitrella patens
 ; US-09-828-303-4

Alignment Scores:	
Pred. No.:	1,586-101
Score:	957.00
Percent Similarity:	99.4%
Best Local Similarity:	99.4%
Length:	1531
Matches:	187
Conservative:	0
Mismatches:	1

Query Match: 97.268 Indels: 1
DB: 3 Gaps: 0
US-10-716-089-20 (1-188) x US-09-828-303-4 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGluProValMet 20
DB 655 ATGGCCACCGAGCGTGTCTCTCAGAGAGACACTCCGAGGCCCTTAGGGCTCAGATTATG 714
QY 21 CysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCysSerLysCys 40
DB 715 TCAGAAGACCTTTGCGGCTTCTTCGAGCAGCACTACCATGGGGGTGTGTGCTGAGATGC 774
QY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGACAGTCATGACGCG-AGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHisGluThr 80
DB 834 GCTCAGGCGACATCTGCCACAGCTGCTGCTTTCAGCCCCCGCTCTCTGACATGAGACC 893
QY 81 LysLeuThrCysGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGCGAGGTGAGAGAACATGATGTGCGCGATCAATCTTCAGCTATCA 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla 120
DB 954 CAAGACCTGGTTAACCCCGCTGACGCTGCCCTCAGCAGTGAAGTCTCTATCCGAGCT 1013
QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGCCCAATGATGCGATCTTGAGAGAGCGTGTGGATTGACAGCA 1073
QY 141 PheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSerAspLysHisThr 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTACTCGGCTTTACATCGGATCGAGCAAAACACT 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal 180
DB 1134 TGCACATATGACTACAAAGCCGAGGGCAGAGCGATGCGAAGCTATCTCTTGTTC 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 4
US-10-716-089-4
; Sequence 4, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/10/716,089
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: 60/196,001
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Physcomitrella patens
US-10-716-089-4

Alignment Scores:
Pred. No.: 1,58e-101 Length: 1531
Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0

Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.268 Indels: 1
DB: 7 Gaps: 0
US-10-716-089-20 (1-188) x US-10-716-089-4 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGluProValMet 20
DB 655 ATGGCCACCGAGCGTGTGTCTCAGAGAGACACTCCGAGGCCCTTAGGGCTCAGATTATG 714
QY 21 CysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCysSerLysCys 40
DB 715 TCAGAAGACCTTTGCGGCTTCTTCGAGCAGCACTACCATGGGGGTGTGTGCTGAGATGC 774
QY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGACAGTCATGACGCG-AGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHisGluThr 80
DB 834 GCTCAGGCGACATCTGCCACAGCTGCTGCTTTCAGCCCCCGCTCTCTGACATGAGACC 893
QY 81 LysLeuThrCysGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGCGAGGTGAGAGAACATGATGTGCGCGATCAATCTTCAGCTATCA 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla 120
DB 954 CAAGACCTGGTTAACCCCGCTGACGCTGCCCTCAGCAGTGAAGTCTCTATCCGAGCT 1013
QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGCCCAATGATGCGATCTTGAGAGAGCGTGTGGATTGACAGCA 1073
QY 141 PheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSerAspLysHisThr 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTACTCGGCTTTACATCGGATCGAGCAAAACACT 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal 180
DB 1134 TGCACATATGACTACAAAGCCGAGGGCAGAGCGATGCGAAGCTATCTCTTGTTC 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 5
US-10-425-114-26146
; Sequence 26146, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack B.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecu
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26146
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB4073-024-B8_FLI
US-10-425-114-26146

Alignment Scores:
Pred. No.: 1,35e-39 Length: 945

Score:	425.50	Matches:	83
Percent Similarity:	61.67%	Conservative:	28
Best Local Similarity:	46.11%	Mismatch:	56
Query Match:	43.24%	Indels:	13
AB:	7	Gaps:	2

US-10-716-089-20 (1-188) x US-10-425-114-26146 (1-945)

Qy	6	GlnGlnLThThrSerGlnAlaProGlnGluProValMetCysLysAsnLeuCysGlyIlePhe	27
Db	112	AAAGAGACCTGGATGCCAGGCACTTGAGGAGCAACCTCTTGATCATCAAACTGGGGCTTC	171
Qy	28	PheGlySerGlnAlaThrMetGlyLeuCysSerLysCysTyArgGlnThrValMetGln	47
Db	172	TTCCGCGAGCGCAGCTCAACATGAACATGTCTCTCAAGTCCCAAGAGAGATGATTAATAG	231
Qy	48	AlaLysMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr	67
Db	232	CAGAGCAGGCGCAAGCTGGCT-----GCTCTCTTATGCAAGCAGCTCAATGGCAAC	285
Qy	68	AlaAlaAlaValGlnProProAlaProValIHisGlnThrLysLeuThrCysGluValGlu	87
Db	286	GATGCTGTCATGGAAACA-----GTTGTTGCT	312
Qy	88	ArgThrMetLLeValProHisGlnSerSerSerTyrgInGlnAspLeuValThrProAla	107
Db	313	GGCAACAAGTGGTCTCCCTGCTCTCCATTCGATGGCAACAATGAACGTGCACCCGCT	372
Qy	108	AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerArgProGluProAsn	127
Db	373	GATGTTGCTGACCTGACGAGGGGGGGGGCGGTGATCTCCAAAGGGAAGGTAGGGCCGAC	432
Qy	128	ArgCysGlySerTybArgLysArgValAlGlyLeuThrGlyPheLysCysAlaArgCysGlyAsn	147
Db	433	CGGTGACGACCTTGGACGAAGAAGGGGTGGACTTCAAGATTCACTGCCGGTGGGGAAC	492
Qy	148	LeuTybCysValLeuLeuHisArgTybSerAspLysHisIstArgCysThrTybAspTybLysAla	167
Db	493	TTGTACGTGCATGCACCGCTACTCTCCGACCAAGCAGCACTGCAGATTGCACTATCCGACT	552
Qy	168	AlaGlyGlnGlnAlaIleAlaLysValAlaAspLeuValAlaAlaGluLysValValLys	187
Db	553	GCCTGATGGATGCCCTTCCCAAGGCTATATCCAGTGGGAAGGGCGACCAAGCTGCACAG	612

RESULT 6

```

? Sequence 71431, Application US/10425115
? Publication No. US20040214272A1
? GENERAL INFORMATION:
? APPLICANT: La Rosa, Thomas J.
? APPLICANT: Kovalic, David K.
? APPLICANT: Zhou, Yihua
? APPLICANT: Cao, Yongwei
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
? TITLE OF INVENTION: Plants
? FILE REFERENCE: 38-21(53222)B
? CURRENT APPLICATION NUMBER: US/10/425,115
? CURRENT FILING DATE: 2003-04-28
? NUMBER OF SEQ ID NOS: 369326

```

Alignment Scores:

Pred. No.:	1.35e-39	947
Score:	425.50	83
Percent Similarity:	61.67%	Conservative: 28
Best local Similarity:	46.11%	Mismatches: 56
Query Match:	43.24%	Indels: 13

DB: 8 Gaps: 2

US-10-716-089-20 (1-188) x US-10-425-115-71431 (1-947)

OY		8	GlnGluThrTrpHisSerGlnAlaProGlnIleProValMetCysTyrAsnL	27
Dd		114	AMGGAGACTGGATGCCAGACCACTTAGGGACCCTTCCTGCATCATATA	173
OY		28	PheGlySerGlnAlaThrMetGlyLeucCysSerLysCysTyrArgGluTr	47
Dd		174	TTCGGCACCGCAGCTACCATGAACAATGTGCTCCAACTGCCAACAGAGA	233
OY		48	AlaLysMetThrAlaLeuAlaGlnGlnIleThrGlnAlaAlaGlnAlaTr	67
Dd		234	CAGAGACAGGCCAAGCTGGCT----GCTTCCTCTATCGACAGCATCG	287
OY		68	AlaAlaAlaValAlaGlnProProAlaProValHisGluThrLysLeuThrC	87
Dd		288	GATGCTGTCATGGAAACA-----	314
OY		88	ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnAsnLeuV	107
Dd		315	GGCAACACAGGTTGCTGCTGCTCCCAATCGAGTTGCAAAACATGAACG	374
OY		108	AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerxrpP	1277
Dd		375	GATGTTGCTGACCTACGAGAGGGCGCGCGTGATCTCCAAAGGGAAAGC	434
OY		128	ArgCysGlySerCysValArgLysArgValGlyLeuThrGlyPheLysCysAl	147
Dd		435	CGGTGCACGACTTGCAGGAAGAAGGGTTGAGATTCAACGATTAACCTGCC	494
OY		148	LeuTyrCysAlaLeuHisArgTyrSerAspLysHisIleThrCysThrTrpAf	1677
Dd		495	TTTGTACTGTGCATGSCAACCGCTACTCCGACCAAGACAGCATGCAAGTTCG	554
OY		168	AlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuValValAlaGluL	187
Dd		555	GCTGTAGGAGATGTCATTCGCCAAGGCTATATCCAGGTGGTGAAGCGCGACN	614

RESULT 7

```

US-10-767-701-15574
? Sequence 15574, Application US/10767701
? Publication No. US20040172684A1
? GENERAL INFORMATION:
? APPLICANT: Kovalic, David K.
? APPLICANT: Zhou, Yihua
? APPLICANT: Cao, Yongwei
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecu
? TITLE OF INVENTION: Plants and Uses Thereof For Plant Impro
? FILE REFERENCE: 38-21 (53535)B
? CURRENT APPLICATION NUMBER: US/10/767,701
? CURRENT FILING DATE: 2004-01-29
? NUMBER OF SEQ ID NOS: 63128
? SEQ ID NO 15574
? LENGTH: 1142
? TYPE: DNA
? ORGANISM: Sorghum bicolor
? FEATURE:
? OTHER INFORMATION: Clone ID: SORBI-28MAV03-CLUS3087_1
US-10-767-701-15574

```

Alignment Scores:

Pred. No.:	1,726-39	Length:	11,142
Score:	425.50	Matches:	83
Percent Similarity:	61.67%	Conservative:	28
Best Local Similarity:	46.11%	Mismatches:	56
Query Match:	43.24%	Indels:	13
DB:	7	Gaps:	2

US-10-716-089-20 (1-188) x US-10-767-701-15574 (1-1142)

QY 8 GlnGluThrThrSerGlnAlaProGlnGlyProValMetCysLysAsnLe

```
DB 112 AAGAGACTGATGATCCAGGACCTGAGGACCCATCTTGCATCAATACTCCGGCTTC 231
QY 28 PheGlySerGlnAlaThrMetGlyLeuCySerIysCySerTyArgGluThrValMetGln 47
DB 232 TTGGCAGCGGACGATGACATGATGCTCTCAAGTCCACAGAGATGATGATTAAGAAG 291
QY 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
DB 232 CAGGAGCAGGCCAAGCTGGCT-----GCCCTCTATGACAGACATCTGTCATGCGAAC 345
QY 68 AlaAlaAlaValAlaGlnProProAlaProValHisGluThrIleuThrCySerGluValGlu 87
DB 346 GATCGCTGTCATGGAACCA-----GTTGTTGCT 372
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyArgGlnAspLeuValThrProAla 107
DB 373 GCGAACACAGTGGTGGCTGCTGCTCCATCGAGTTGCACAAACATGAACGTGACGCCGCT 432
QY 108 AlaAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluProAsn 127
DB 433 GATGTTGCTGACCTAGCCAGGGGGCGGGGTGATCTCCAAAGGAGAGGTAGCGCGAAC 492
QY 128 ArgCySerGlySerCyArgIysArgValGlyLeuThrGlyPheIysCySerGlyAsn 147
DB 493 CGGTCCAGACACTTGCAGAGAGGGGTGGACTTACAGATTCAACTGCCGGGTGGAGAC 552
QY 148 LeuTyTCyValAlaLeuHisArgTySerAspIysHisThrCySerThrAspTyArgVala 167
DB 553 TTGTACTGGGACTGCACCGGCTACTCCGACAGACAGACTCGAAGTTGCACTATCGGACT 612
QY 168 AlaGlyGlnGluAlaIleAlaIysAlaAsnProLeuValValAlaGluIysValIys 187
DB 613 GCTGCTAGGAGATGCTATTGCGCAAGGCTAATCCAGTGGTAGGCGGACAGACTCGACAAAG 672

RESULT 8
US-10-437-963-21113
/ Sequence 21113, Application US/10437963
/ Publication NO. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 21113
/ LENGTH: 2671
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_26415C.1
US-10-437-963-21113
```

```
Alignment Scores:
Pred. No.: 2,586-38 Length: 2671
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 7 Gaps: 2
```

US-10-716-089-20 (1-188) x US-10-437-963-21113 (1-2671)

QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCyIysAsnLeuCySerIysPhe 27

```
DB 1457 AAGAGACTGATGATCCAGGACCGCAGAGGCCCGGATCTTGCATCAATA 1516
QY 28 PheGlySerGlnAlaThrMetGlyLeuCySerIysCySerTyArgGluThr 47
DB 1517 TTGGCAGTGGCGGATGACATGATGCTCTCAAGTCCACAGAGATGATGATTAAGAAG 1576
QY 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThr 67
DB 1577 CAGGAGCAGGCCAAGCTGGCA-----GCCCTCTATGACAGACATGCTG 1630
QY 68 AlaAlaAlaValAlaGlnProProAlaProValHisGluThrIleuThrCySer 87
DB 1631 GATTCGGGAGAGAACCAATTATGCTGCTGCTCACGCTGAAGTAGCTGTTG 1690
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyArgGlnAspLeuValThrProAla 107
DB 1691 GTGAAGACGCTGTT-----G 1717
QY 108 AlaAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluProAsn 127
DB 1718 GAGATTGCTGATCTTACGAGGGGGGTGACGCTGTAACCCCAAGGGAGGAGG 1777
QY 128 ArgCySerGlySerCyArgIysArgValGlyLeuThrGlyPheIysCySer 147
DB 1778 CGGTCCAGACACTTGCAGAGAGGGGTGGACTTACAGATTCAACTGCCGCTG 1837
QY 148 LeuTyTCyValAlaLeuHisArgTySerAspIysHisThrCySerThrTyArg 167
DB 1838 TTGTACTGGGACTGCACCGGCTACTCCGACAGACAGACTCGAAGTTGCACTATCGGACTTC 1897
QY 168 AlaGlyGlnGluAlaIleAlaIysAlaAsnProLeuValValAlaGluI 187
DB 1898 GCTGCTAGGAGATGCTATTGCGCAAGGCTAATCCGCTGTTGAGGCGGAGAG 1957

RESULT 9
US-10-425-114-27574
/ Sequence 27574, Application US/10425114
/ Publication NO. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Liu, Jindong
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack B
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Tabaska, Jack B
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
/ FILE REFERENCE: 38-21(53131)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 27574
/ LENGTH: 1032
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4732-053-A11_FLI
US-10-425-114-27574
```

```
Alignment Scores:
Pred. No.: 1,716-38 Length: 1032
Score: 416.50 Matches: 85
Percent Similarity: 62.09% Conservative: 28
Best Local Similarity: 46.70% Mismatches: 56
Query Match: 42.33% Indels: 13
DB: 7 Gaps: 4
```

US-10-716-089-20 (1-188) x US-10-425-114-27574 (1-1032)

QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCyIysAsnLeu 27

DB 134 AAGAGCTGGGTGCGCAGGCCCGGAGGACCCATCTCTGCAATCAATA 193

Qy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIysCySerTyraArgGluThrValMetCln 47
Db 194 TTTCGACGGCGGACCATGAACATGTGCTTAAGTCCACAGAGATGATTAACGAG 253
Qy 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
Db 254 CAGATCAAGCGCAAGTGGCT-----GCTTCCTTATTCACAGCATGTGTAACGGCAGC 307
Qy 68 AlaAlaAlaValGlnProProAlaProValHisGluThrLysLeuThrCysGluValGlu 87
Db 308 GAGCGCCGTCAGAGCCCGGTGTGTGCTGTCGACAGCAACAG-----GTACTA 352
Qy 88 ArgThrMetIleValProHisGlnSerSerSerTyraGlnInaPheLeuValThrProAla 107
Db 353 CGTGTGCCCAAGTCAGATTGCAACATGAACGTGCACAG-----CCCCCT 400
Qy 108 -----AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerArgProGlu 125
Db 401 GATGTTCGCAACCCAGAGAGGGGGTGGCGCGCGATCTCCAAAGGGGGAGGTAGCG 460
Qy 126 ProAlaArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysArgCys 145
Db 461 CCGAACCCGTCAGCGCTGCAGAGAGGGGTGGCTCACCGGAGATTCACCTGCCGGTGT 520
Qy 146 GlyAlaLeuTyraValAlaLeuHisArgTyraSerAspLysHisThrCysThrTyraAspTyr 165
Db 521 GGGAACTGTGACTGGCGCTCCACCGCTACTCCGAAAGCAGACTGCAAGTTGCACCTAC 580
Qy 166 LysAlaAlaGlyGlnGlnAlaIleAlaLysAlaAsnProLeuValAlaGluLysVal 185
Db 581 CGGACTGTGCACAGGAGCGCATTTGCCAAGGCTATCCGGTGTGAAGCAGACAAAGCTC 640
Qy 186 ValLys 187
Db 641 GACAAG 646

RESULT 10
US-10-437-963-37809/c
/ Sequence 37809, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 37809
/ LENGTH: 1171
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_41503C.1
US-10-437-963-37809

Alignment Scores:
Pred. No.: 2,01e-38 Length: 1171
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47
Query Match: 42.33% Indels: 35
DB: 7 Gaps: 3
US-10-716-089-20 (1-188) x US-10-437-963-37809 (1-1171)

Qy 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysLysAsnL 27
Db 890 AAGAGGCTGTGCTCCACAGACCCGAGAGGCCCAATCTTTCATCAATAT 831
Qy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIysCySerTyraArgGlu 47
Db 830 TTTCGACGGCGGACCATGAACATGTGCTTCAAGTCCACAGAGAGA 771
Qy 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
Db 770 GAGAGCAGGCCAAGCTTGCTGCTCCTCCATGCATAGCATGTCAATG 711
Qy 61 -----AlaGlnAlaThrSerAlaThrAlaAlaValG 75
Db 710 GGGAAAGACATATTGTTCGTCACAGTGTGACACGGCGGTGGCTG- 663
Qy 76 ProValHisGluThrLysLeuThrCysGluValGluArgThrMetIle 95
Db 662 -----GTCAGGTTCGAGCGGAAGACGCTCG 636
Qy 96 SerSerSerTyraGlnInaPheLeuValThrProAlaAlaAlaProG 115
Db 635 -----GTGCAGCTTACCGATGTCCGGGCA 600
Qy 116 SerSerIleAlaAlaProSerArgProGluProAlaArgCysGlySerC 135
Db 599 GTTGCTGTATCCCAAGCTGCAAGAGAGGCCGAACCGGTGCGTCACTC 540
Qy 136 ValGlyLeuThrGlyPheLysCysArgCysGlyAlaLeuTyraValLe 155
Db 539 GTTGAGGTGACGCGGATTCACCTGCCGGTGGTAAACATGATCTGGCGCT 480
Qy 156 SerAspLysHisThrCysThrTyraAspTyrLysAlaAlaGlyGlnAla 175
Db 479 TCCGACCAAGCATGAATGCAGTTCGACTACCGGACTGCGGCTAGGGATG 420
Qy 176 AlaAsnProLeuValAlaAlaGluLysValLys 187
Db 419 GCCAACCCAGTGTGAAGCTGAGAGACTCCAGCAAG 384

RESULT 11
US-10-437-963-37808/c
/ Sequence 37808, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Mo
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 37808
/ LENGTH: 1249
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_41502C.1
US-10-437-963-37808

Alignment Scores:
Pred. No.: 2.18e-38 Length: 1249
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47

Query Match: 42.33% Indels: 35
DB: 7 Gaps: 3
US-10-716-089-20 (1-188) x US-10-437-963-37808 (1-1249)

QY 8 GInGluThrThrSerGlnAlaProGluGlyProValMetCysValAsnLeuCysGlyPhe 27
DB 513 AAGAGGCTGGCTGCCAGCAGCCGAGGCGCCATCTTGCAATCAATACCGGCTTC 454
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerLeuCysTyrArgGluThrValMetGln 47
DB 453 TTGGCAGCGCGCGCAGCATGATGCTCTCCAGTGCACAGAGATGATCATGAG 394
QY 48 AlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
DB 393 GAGGACGAGCCCAAGTGTCTGCTCTCCATCGATATGATGTCATGTTGTGACGCT 334
QY 61 -----AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAla 75
DB 333 GGGAGGACATATTGTTGCTGCTGCAATGAGTACGCGGCTGCGCTG----- 286
QY 76 ProValHisGluThrLysLeuThrCysGluValGluArgThrMetLeValProHisGln 95
DB 285 -----GCTCAAGTCAAGCGGAGACGCTCGT----- 259
QY 96 SerSerSerTyrGlnGlnIleuLeuValThrProAlaAlaAlaProGlnAlaValLys 115
DB 258 -----GTGCAAGCTCAACGATGTCGCGGCAACGCGCTCGT----- 223
QY 116 SerSerIleAlaAlaProSerArgProGluProAsnArgCysGlySerCysArgLysArg 135
DB 222 GTTGTCTGTATGCCCAAGCTCAAGAGGCGCCAGCGGCTGCTACCTGTGAGAGAG 163
QY 136 ValGlyLeuThrGlyPheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyr 155
DB 162 GTTGGGCTCAGCGGATTCATGCGCGTGCAGTACATGATGCTGCTGCAACGCTAC 103
QY 156 SerArgLysHisThrCysThrTyrArgTyrLysAlaAlaGlyGlnGluAlaIleAlaLys 175
DB 102 TCCGACACCATGAAAGCCAGTTCGACTACCGACGCTGAGGATCCATCGCCAG 43
QY 176 AlaAsnProLeuValAlaGluLysValLys 187
DB 42 GCCAACCCAGTGTGAGGCTGAGAGCTGCACAG 7

RESULT 12
US-10-425-114-22047
/ Sequence 22047, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 22047
/ LENGTH: 736
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3356-023-Fl1_FLI
US-10-425-114-22047

Alignment Scores:
Pred. No.: 1.65e-38 Length: 736
Score: 415.00 Matches: 85

Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 7 Gaps: 4
US-10-716-089-20 (1-188) x US-10-425-114-22047 (1-736)

QY 8 GInGluThrThrSerGlnAlaProGluGlyProValMetCysValAsnLeu 27
DB 113 AAGAGGCTGGCTGCCAGCAGCCGAGGCGCCATCTTGCAATCAATACCGGCTTC 172
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerLeuCysTyrArgGluThrValMetGln 47
DB 173 TTGGCAGCGCGCGCAGCATGATGCTCTCCAGTGCACAGAGATGATCATGAG 232
QY 48 AlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAla 67
DB 233 CAGGATCAAGCCCAAGTGTCT-----GCTCTCTATGACAGCATC 286
QY 68 AlaAlaAlaValGlnProProAlaProValHisGluThrLysLeuThr 87
DB 287 GAGCGCTCATGAGCGCGTGTGCTGCTGCGCAGCAACAG----- 331
QY 88 ArgThrMetIleValProHisSerSerSerTyrGlnGlnAsnLeu 107
DB 332 GCTGTGCCCAAGTTCAGTTCGAAACATGAACTGTCAGCAG----- 379
QY 108 ---AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSer 126
DB 380 GATGTGCCGAGCCAGCAGGAGGCGTGCAGATCTCCAAAGGCGGGA 439
QY 127 AsnArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLys 146
DB 440 AACCGCTGACGCGCTCGCAGAGAGGCGTTCGACTTACGCGGATTCAC 499
QY 147 AsnLeuTyrCysAlaLeuHisArgTyrSerArgLysHisThrCysThr 166
DB 500 AACTGTACTGCGCACTCCACCGCTATCCGACAGCAGACTGCAAGT 559
QY 167 AlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuValAla 186
DB 560 ACTGTGCCAGGAGCCATTCGCAAGGCTTAATCCGGTGTGAAGCAG 619
QY 187 Lys 187
DB 620 AAG 622

RESULT 13
US-10-425-114-28524
/ Sequence 28524, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 28524
/ LENGTH: 776
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4758-034-E9_FLI
US-10-425-114-28524

Alignment Scores:
Pred. No.: 1.65e-38 Length: 736
Score: 415.00 Matches: 85

Pred. No.: 1.77e-38 Length: 776
Score: 415.00 Matches: 85
Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 7 Gaps: 4

US-10-716-089-20 (1-188) x US-10-425-114-28524 (1-776)

```
QY      8 GlnGluThrThrSerGlnAlaProGluGluGlyProValMetCysGlyAsnLeuCysGlyPhe 27
      133 AAGAGGCTGGGTGCTCAGAGCCCGCAGAGGACCCATCTCTGCATCAATAACTGTGGCTTC 192
QY      28 PheGlySerGlnAlaThrMetGlyLeuCysSerIysCysGlyTyrArgGluThrValMetGln 47
      193 TTCGACAGCGCGGCACCATGACATGTGCTTTAGTGCCACAGAGATGATTACGAAG 252
QY      48 AlalyMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
      253 CAGGATCAGGCCAAGCTGGCT-----GCTCTCTATCGACAGCATGTGTAAAGCGCAGC 306
QY      68 AlAlaAlaValAlaGlnProProAlaProValHisGluThrIysLeuThrCysGluValGlu 87
      307 GACGCCGTATGAGACCGGTTGCTTGTCTGGCAGCAACAG-----GTAGTA 351
QY      88 ArgThrMetIleValProHisGlnSerSerSerTyrGlnGlnAspLeuValThrProAla 107
      352 GCTGTGCCCAAGTGCAGATTGCACAAATGACGTGCACAG-----CCCGCT 399
QY      108 --AlAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluPro 126
      400 GATGTGGCCGACCCAGAGAGGGGTGGCGGCATCTCCAAAGGGGAGTAGGGCGC 459
QY      127 AsnArgCysGlySerCysArgIysArgValGlyLeuThrIysPheIysCysArgCysGly 146
      460 AACCGGTCCAGCGCTGCAGAGAGGGTGGCTCAGCGGATTCACTGCGCGGTGGG 519
QY      147 AsnLeuTyrCysAlaLeuHisArgTyrSerAspIysHisIleThrCysThrTyrAspTyrIys 166
      520 AACTGTGACTGCGCCCTCCACCGCTACTCCGACAGCAGACTGCAGATTGACTACCGG 579
QY      167 AlAlaGlyGlnGlnAlaIleAlaIysAlaAsnProLeuValValaGluIysValVal 186
      580 ACTGCTGCAGAGGAGCGCATTCGCAAGGCTAATCCGGTGTGAAGCAGACAAAGCTCGAC 639
QY      187 Lys 187
      640 AAG 642
DB
```

RESULT 14

```
US-10-425-114-22082
; Sequence 22082, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22082
; LENGTH: 821
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-040-B2_FLI
US-10-425-114-22082
```

Alignment Scores:
Pred. No.: 1.9e-38 Length: 821
Score: 415.00 Matches: 85
Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 7 Gaps: 4

US-10-716-089-20 (1-188) x US-10-425-114-22082 (1-821)

```
QY      8 GlnGluThrThrSerGlnAlaProGluGluGlyProValMetCysGlyAsnLeu 27
      113 AAGAGGCTGGGTGCTCAGAGCCCGCAGAGGACCCATCTCTGCATCAATAACTGTGGCTTC 172
QY      28 PheGlySerGlnAlaThrMetGlyLeuCysSerIysCysGlyTyrArgGluThrValMetGln 47
      173 TTCGACAGCGCGGCACCATGACATGTGCTTTAGTGCCACAGAGATGATTACGAAG 232
QY      48 AlalyMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
      233 CAGGATCAGGCCAAGCTGGCT-----GCTCTCTATCGACAGCATGTGTAAAGCGCAGC 286
QY      68 AlAlaAlaValAlaGlnProProAlaProValHisGluThrIysLeuThrCys 87
      287 GACGCCGTATGAGACCGGTTGCTTGTCTGGCAGCAACAG-----GTAGTA 331
QY      88 ArgThrMetIleValProHisGlnSerSerSerTyrGlnGlnAspLeuValThrProAla 107
      332 GCTGTGCCCAAGTGCAGATTGCACAAATGACGTGCACAG-----CCCGCT 379
QY      108 --AlAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluPro 126
      380 GATGTGGCCGACCCAGAGAGGGGTGGCGGCATCTCCAAAGGGGAGTAGGGCGC 439
QY      127 AsnArgCysGlySerCysArgIysArgValGlyLeuThrIysPheIysCysArgCysGly 146
      440 AACCGGTCCAGCGCTGCAGAGAGGGTGGCTCAGCGGATTCACTGCGCGGTGGG 499
QY      147 AsnLeuTyrCysAlaLeuHisArgTyrSerAspIysHisIleThrCysThrTyrAspTyrIys 166
      500 AACTGTGACTGCGCCCTCCACCGCTACTCCGACAGCAGACTGCAGATTGACTACCGG 559
QY      167 AlAlaGlyGlnGlnAlaIleAlaIysAlaAsnProLeuValValaGluIysValVal 186
      560 ACTGCTGCAGAGGAGCGCATTCGCAAGGCTAATCCGGTGTGAAGCAGACAAAGCTCGAC 619
QY      187 Lys 187
      620 AAG 622
DB
```

RESULT 15

```
US-10-425-114-21492
; Sequence 21492, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 21492
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-040-B2_FLI
US-10-425-114-21492
```

! OTHER INFORMATION: Clone ID: L183354-001-F10_FLI
US-10-425-114-21492

Alignment Scores:
Pred. No.: 1.92e-38 Length: 826
Score: 415.00 Matches: 85
Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 7 Gaps: 4

US-10-716-089-20 (1-188) x US-10-425-114-21492 (1-826)

```
QY      8 GInGluThrThrSerGlnAlaProGluGlnProValMetCysGlyAsnLeuCysGlyPhe 27
      8 AAGGAGGCTGGGTGGCCAGGCCCCGAGGAGCCATCTCTGCATCAATACTGTGGCTTC 67
QY      28 PheGlySerGlnAlaThrMetGlyLeuCysSerGlyCysTyrArgGluThrValMetGln 47
      68 TTCGGCAGCCCGCCGACCATGACATGTGCTTAAGTGCACAGAGATGATACGAG 127
QY      48 AlaIyMetThrAlaLeuAlaGluGlnAlaThrGlnAlaIaGlnAlaThrSerAlaThr 67
      128 CAGGATCAGGCCAAGCTGGCT-----GCCTCTATGCACAGCATCGTGAACGGCAGC 181
QY      68 AlaAlaAlaValGlnProProAlaProValHisGluThrIlyLeuThrCysGluValGlu 87
      182 GACGCCGTATGGAGCCGCTGTGGTGTGGCAGCACACG-----GTAGTA 226
QY      88 ArgThrMetIleValProHisGlnSerSerTyrGlnGlnAspLeuValThrProAla 107
      227 GCTGTGGCCCAAGTCCAGTTGCCAAACATGAACTGCAGCAG-----CCGCT 274
QY      108 ---AlaAlaAlaProGlnAlaValIlySerSerIleAlaAlaProSerArgProGluPro 126
      275 GATGTTGCCGAGCCAGCGAGGGGGTGGCGCATCTCCAAAGGGGGAAGTAAAGGCCG 334
QY      127 AsnArgCysGlySerCysArgGlyValGlyLeuThrGlyPheIlyCysArgCysGly 146
      335 AACCGGTGACGCCCTGCAGGAGAGGGGTGGCTCACGGATTCAACTGCCGCTGGG 394
QY      147 AsnLeuTyrCysAlaLeuHisArgTyrSerAspIlyHisThrCysThrTyrAspTyrIly 166
      395 AACTTTACTGCGCGCTCCACCGCTACTCCGACACAGCAGACTGCAGATTCACTACCG 454
QY      167 AlaAlaGlyGlnGlnAlaIleAlaIleAlaIleAsnProLeuValAlaIaGluIlyVal 186
      455 ACTGCTGCCAGGAGCGCATTTGCCAAGGCTAATCCGTTGTGAAGGCAGCAAGCTCGAC 514
QY      187 Ily 187
DB      515 AAG 517
```

Search completed: December 8, 2005, 18:57:54
Job time : 795 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:44:44 ; Search time 249 Seconds

(without alignments)
282.268 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSEGTSGARDEGPM.....GGEALAKANPLVAKVKYKF 188

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cg2_1/USPTO_spool_p/US10716089/rnatc 05122005 094345 13229/app_query.fasta_1.327
-DB=Published Applications NA New -OPMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOCL=0 -LOOPEXT=0 -UNITS=hits -START=1 -END=-1 -MATRIX=blowsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10716089_@CGN 1.1 184_@rnatc 05122005 094345_13229
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOF=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New.*

- 1: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 2: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 6: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 7: /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 8: /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
- 9: /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
- 10: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189.5	19.3	1067	US-10-750-185-47138	Sequence 47138, A
2	180	18.3	1563	US-10-750-185-52649	Sequence 52649, A
3	90.5	9.2	864	US-10-467-657-1869	Sequence 1869, Ap
4	88.5	9.0	2133	US-11-045-802-10	Sequence 10, Appl
5	86.5	8.8	3950	US-10-821-234-253	Sequence 253, App
6	86	8.7	2089	US-10-750-185-60675	Sequence 60675, A
7	86	8.7	3513	US-10-467-657-7841	Sequence 7841, Ap
8	86	8.7	7893	US-11-186-731-3	Sequence 3, Appl

C 9	86	8.7	8106	7	US-11-186-731-1	Seq	p1i
C 10	86	8.7	23907	7	US-11-186-731-6	Seq	p1i
C 11	86	8.7	24120	7	US-11-186-731-4	Seq	p1i
C 12	85	8.6	1323	7	US-11-087-100-21	Seq	p1i
C 13	85	8.6	6177	7	US-11-087-100-3	Seq	p1i
C 14	83	8.4	2815	6	US-10-821-234-165	Seq	p1i
C 15	82.5	8.4	801	6	US-10-508-263-103	Seq	App
C 16	82.5	8.4	1076	6	US-11-102-240-45	Seq	App
C 17	82.5	8.4	2159	6	US-10-131-826A-87	Seq	p1i
C 18	82	8.3	1161	7	US-10-858-730-170	Seq	p1i
C 19	81	8.2	16816	7	US-11-121-086-3	Seq	p1i
C 20	80.5	8.2	1991	6	US-10-750-185-57360	Seq	'A
C 21	80.5	8.2	2825	6	US-10-750-185-56778	Seq	'A
C 22	80.5	8.2	3513	6	US-10-467-657-7841	Seq	Ap
C 23	80	8.1	1022	6	US-10-750-185-65872	Seq	'A
C 24	80	8.1	8424	6	US-10-821-234-47	Seq	p1i
C 25	80	8.1	10373	6	US-10-821-234-64	Seq	p1i
C 26	79.5	8.1	2133	7	US-11-045-802-11	Seq	p1i
C 27	79.5	8.1	2302	7	US-11-110-082-20	Seq	p1i
C 28	79	8.0	1102	6	US-10-131-826A-299	Seq	App
C 29	79	8.0	1392	7	US-11-182-946-1	Seq	App
C 30	79	8.0	1511	6	US-10-750-185-37762	Seq	'A
C 31	79	8.0	6058	6	US-10-770-726-17	Seq	p2i
C 32	79	8.0	168516	7	US-11-121-086-3	Seq	p1i
C 33	79	8.0	169495	7	US-11-121-086-61	Seq	p1i
C 34	78.5	8.0	1017	6	US-10-909-125-823	Seq	App
C 35	78.5	8.0	3334	6	US-10-821-234-152	Seq	App
C 36	78.5	8.0	8424	6	US-10-821-234-47	Seq	p1i
C 37	78.5	8.0	14342	6	US-10-821-234-303	Seq	App
C 38	78.5	8.0	78869	7	US-11-075-185-1	Seq	p1i
C 39	78.5	8.0	197096	7	US-11-121-086-107	Seq	App
C 40	78	7.9	407	7	US-11-108-172-174	Seq	App
C 41	78	7.9	1140	6	US-10-793-626-2809	Seq	App
C 42	78	7.9	1467	6	US-10-821-234-802	Seq	App
C 43	78	7.9	1675	6	US-10-750-185-41941	Seq	'A
C 44	78	7.9	2260	7	US-11-186-284-120	Seq	App
C 45	78	7.9	2988	6	US-10-793-626-4132	Seq	Ap

ALIGNMENTS

RESULT 1
US-10-750-185-47138/C
Sequence 47138, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 47138
LENGTH: 1067
TYPE: DNA
ORGANISM: Bovine 19866881619274
US-10-750-185-47138
Alignment Scores:
Pred. No.: 7.86e-09
Score: 189.50
Percent Similarity: 52.24%
Best Local Similarity: 32.84%
Query Match: 19.26%
Length: 1067
Matches: 66
Conservative: 39
Mismatch: 61
Indels: 38

DB: 6 Gaps: 8
US-10-716-089-20 (1-188) x US-10-750-185-47138 (1-1067)
QY 18 ProValMetCysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCys 37
1034 CCTGTGCTTGTTCGTGCGCTGACAGATTATTATGAAACCCATAGACGAATGGCATGTGT 975
QY 38 Ser-----LysCysTyrArgGluThrValMetGln----- 47
974 TCAGTCCGCAATAGAAATCATCTTCAAAATAGTTCAGAAATAGATGTAGAAATAGC 915
QY 48 -----AlaLysMetThrAlaLeuAlaGlu-----GlnAlaThr----- 58
914 CCCCTGTACTTCTGTCTGTCTGTCTGTCTGTCTTTCACGCCACAGACCAAGCCG 855
QY 59 -----GlnAlaAlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAla 75
854 AGTGTCCAGAAAGCTCAGTCAAGCGCTAGACGCTTCGTTCATGATGACCAAGCCCT 795
QY 76 ProValIHisGluThrLysLeuThrCysGluValGluArgThrMetIleValProHisGln 95
794 GTATCACAATCACTCACTTATCA-----GAATCTGTAGTGTCTTCCCAAGTGA 744
QY 96 SerSerSerTyrGlnGlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLys 115
743 CAGTACATCTATGACAAAGCAATCTGGAACAGATGACCTGCAAACTGCAGATACGA 684
QY 116 SerSerIleAlaAlaProSerArgProGluPro-----AsnArg--- 128
683 AACAGATAGCAGCCGCTGGAAGAGA-AAACCAAGTCTTGAAAAACCAAAACGAAAA 625
QY 129 -----CysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysArgCysGly 146
624 AGAATCTGTTTATGTGCGAAGAAAGTGGAGCTTG---TTTGAAGCCAGGT-6GA 570
QY 147 AsnLeuTyrCysAlaLeuHisArgTyrSerAspLysHisIleThrCysThrTyrAspTyrLys 166
569 AATGTTTACTGTAGGTATACCATTTACTCATATGATGACACACTGCTCTTACCAATTAAAGA 510
QY 167 AlaAlaGlyGlnGlnAlaIleAlaLysAlaAsnProLeuValAlaGluLysValVal 186
509 CTGATGGCTGCTGAGAAATCAGAAAAAATCCAGTACTTGTTGGTGAAGATCCAG 450
QY 187 Lys 187
DB 449 MAG 447
RESULT 2
US-10-750-185-52649/C
; Sequence 52649, Application US/10750185
; Publication No. US2005026063A1
; GENERAL INFORMATION:
; APPLICANT: NMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52649
; LENGTH: 1563
; TYPE: DNA
; ORGANISM: Bovine 19866880938011
US-10-750-185-52649

Alignment Scores:
Pred. No.: 1.04e-07 Length: 1563
Score: 180.00 Matches: 33
Percent Similarity: 75.00% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 11
Query Match: 18.29% Indels: 1
DB: 6 Gaps: 0
US-10-716-089-20 (1-188) x US-10-750-185-52649 (1-1563)
QY 139 ThrGlyPheLysCys-ArgCysGlyAsnLeuTyr-CysAlaLeuHisArg 158
927 ACAGGCTTTCAGCTCCAGATGTGGAATTTCTTTGTGTGACTTCACCGT 868
QY 158 HisIleThrCysThrTyrAspTyrLysAlaAlaGlyGlnGlnAlaIleAla 178
867 GCACAACGTGTCCATATGATTCAAAGCAGAACTGCAGCAAAAATCAGAG 808
QY 178 OleuValValAlaGluLysVal 185
DB 807 AGTTGTGTGGCTGAAAAAATT 786
RESULT 3
US-10-467-657-1869/C
; Sequence 1869, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASTIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1869
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1869
Alignment Scores:
Pred. No.: 21.4 Length: 864
Score: 90.50 Matches: 41
Percent Similarity: 40.14% Conservative: 16
Best Local Similarity: 28.87% Mismatches: 49
Query Match: 9.20% Indels: 36
DB: 6 Gaps: 7
US-10-716-089-20 (1-188) x US-10-467-657-1869 (1-864)
QY 32 AlaThrMetGlyLeuCysSerLysCysTyrArgGluThrValMetGlnAla 51
DB 371 GCGGCATCGGCAGGTGTTCT-----TCAA 336
QY 52 AlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr- 71
DB 335 AATTCACCGCAAGAGGCAACCGTGCACTTACCGCTGAAATCGGAGCTT 276
QY 71 GlnProProAlaProAla-----HisGluThrLysLeuThr 87
DB 275 TGTTCACACACACCGGACGCGATTTCGGTATATAATCCCGTAAAG 219
QY 87 LysThrMetIleValAlaProHisGln-----SerSerSerTyrGln 104
DB 218 TCGGCATTCGCGCGTGCAGACGACCGACCAATCAAAAATATACACGA 159

Pred. No.: 900 Length: 7893
 Score: 86.00 Matches: 37
 Percent Similarity: 37.40% Conservative: 12
 Best Local Similarity: 28.24% Mismatches: 71
 Query Match: 8.74% Indels: 11
 DB: 7 Gaps: 3

US-10-716-089-20 (1-188) x US-11-186-731-3 (1-7893)

QY 29 GYSGRGLAIAIATHMeGlyLeuCySerLySCTYrArgGluThrValMeGlnAla 48
 DB 2183 GGGTCTTCATGAGATGTGGGGGTCCACCTGCATGCTTCCATCTTGTACAG 2124
 QY 49 LyMeCTHrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThrAla 68
 DB 2123 CTGATGACAGCTTGGGTGTGCCCGCTGACGGAGCCAGCTTGATCTGTCCACCCAG 2064
 QY 69 AAlaAlaValGlnPro-ProAlaProValHleGluThrLySLeuThrCyGluValGluAr 88
 DB 2063 TCGGCTGTGCATGCGGCCAGCTCTCTTCAAAAGTCCGGGGGCCGCACACAGCGGCC 2004
 QY 88 GThMeTleValProHleGlnSerSerSerTyrglnGlnAep-----LeuValTh 105
 DB 2003 AGACGCTGTGGATGCCACAGATCTCTTCAACCACAGACTCTGTATATGGCTGTCCGT 1944
 QY 105 rProAlaAlaAlaAlaProGlnAlaValLySerSerTleAlaAlaProSerArg----- 123
 DB 1943 GCTTCAGACAGTACTTGGCCACCGAGTCTCCCTCTCTCCACACCTCGAAGGCGCG 1884
 QY 124 -----ProGluProAlaArgCyGlySerCySArgLySArgValGlyLeuThrGlyPh 141
 DB 1883 TCATCCCCCTCCACCTGCTGTCTGATGATGCTGCTCAAGCTTCAATCATGTTCGGAGAC 1824
 QY 141 eLyCyS-----ArgCyGlyAsnLeu 148
 DB 1823 ACGTAGCTGACGATACGGTGCAGGAGTCTC 1793

RESULT 9

US-11-186-731-1/c
 / Sequence 1, Application US/11186731
 / Publication No. US20050255521A1
 / GENERAL INFORMATION:
 / APPLICANT: Kapeller-Libermann, Rosana
 / APPLICANT: Acton, Susan L.
 / TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
 / FILE REFERENCE: MP12001-047PIRCP1(M)
 / CURRENT APPLICATION NUMBER: US/11/186,731
 / PRIOR FILING DATE: 2005-07-21
 / PRIOR APPLICATION NUMBER: US/10/077,130
 / PRIOR FILING DATE: 2002-02-15
 / PRIOR APPLICATION NUMBER: 60/269201
 / NUMBER OF SEQ ID NOS: 9
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 1
 / LENGTH: 8106
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: 5'UTR
 / LOCATION: (1)...(71)
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (72)...(7964)
 / FEATURE:
 / NAME/KEY: 3'UTR
 / LOCATION: (7965)...(8106)
 / US-11-186-731-1

Alignment Scores:
 Pred. No.: 930
 Score: 86.00

Length: 8106
 Matches: 37

Percent Similarity: 37.40% Conservative: 12
 Best Local Similarity: 28.24% Mismatches: 71
 Query Match: 8.74% Indels: 11
 DB: 7 Gaps: 3

US-10-716-089-20 (1-188) x US-11-186-731-1 (1-8106)

QY 29 GYSGRGLAIAIATHMeGlyLeuCySerLySCTYrArgGluThrV 48
 DB 2254 GGGTCTTCATGAGATGTGGGGGTCCACCTGCATGCTTCCATCTTGTACAG 2195
 QY 49 LyMeCTHrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThrAla 68
 DB 2194 CTGATGACAGCTTGGGTGTGCCCGCTGACGGAGCCAGCTTGATCTGTCCACCCAG 2135
 QY 69 AAlaAlaValGlnPro-ProAlaProValHleGluThrLySLeuThrCyG 88
 DB 2134 TCGGCTGTGCATGCGGCCAGCTCTCTTCAAAAGTCCGGGGGCCGCAC/ 2075
 QY 88 GThMeTleValProHleGlnSerSerSerTyrglnGlnAep----- 105
 DB 2074 AGACGCTGTGGATGCCACAGATCTCTTCAACCACAGACTCTTGTATA 2015
 QY 105 rProAlaAlaAlaAlaProGlnAlaValLySerSerTleAlaAlaPro 123
 DB 2014 GCTTCAGACAGTACTTGGCCACCGAGTCTCCCTCTCTCCACACCT 1955
 QY 124 -----ProGluProAlaArgCyGlySerCySArgLySArgValGly 141
 DB 1954 TCATCCCCCTCCACCTGCTGTCTGATGATGCTGCTCAAGCTTCAATCATG 1895
 QY 141 eLyCyS-----ArgCyGlyAsnLeu 148
 DB 1894 ACGTAGCTGACGATACGGTGCAGGAGTCTC 1864

RESULT 10

US-11-186-731-6/c
 / Sequence 6, Application US/11186731
 / Publication No. US20050255521A1
 / GENERAL INFORMATION:
 / APPLICANT: Kapeller-Libermann, Rosana
 / APPLICANT: Acton, Susan L.
 / TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
 / FILE REFERENCE: MP12001-047PIRCP1(M)
 / CURRENT APPLICATION NUMBER: US/11/186,731
 / PRIOR FILING DATE: 2005-07-21
 / PRIOR APPLICATION NUMBER: US/10/077,130
 / PRIOR FILING DATE: 2002-02-15
 / PRIOR APPLICATION NUMBER: 60/269201
 / NUMBER OF SEQ ID NOS: 9
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 6
 / LENGTH: 23907
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-11-186-731-6

Alignment Scores:
 Pred. No.: 3,556+03 Length: 23907
 Score: 86.00 Matches: 37
 Percent Similarity: 37.40% Conservative: 12
 Best Local Similarity: 28.24% Mismatches: 71
 Query Match: 8.74% Indels: 11
 DB: 7 Gaps: 3

US-10-716-089-20 (1-188) x US-11-186-731-6 (1-23907)

QY 29 GYSGRGLAIAIATHMeGlyLeuCySerLySCTYrArgGluThrV 48
 DB 18197 GGGTCTTCATGAGATGTGGGGGTCCACCTGCATGCTTCCAT 18136

Oy		49	lymetherhlaleumlaagluglnalathglnalaalaglmlahphserlathrla	68
Db	18137	CTGATGACAGCGTTGGGTGGTGTGCCCGTCACCGCCAGCGCAGCGTTGATCTGCACCAC	18078	
Oy		69	AlaIalValGlInPro-ProAlaProValHisIuThrLyLeuthrCyGluValGluar	88
Db	18077	TCCGCTTGCAATGTCGGCCAGACTCTCTTCAAAGTCCGGGGGCCGACACAGCAGGGCC	18018	
Oy		88	grhmetilevalprohisginserSerSerrTyrgInIasp-----leuValrh	105
Db	18017	AGACGCTGCGAGATGCCACAGATCTCTTACCCAGAGCTGTGAATGGCTGCGT	17958	
Oy		105	rProAlaIalAalAalProGlnIalValLysSerSerIealAalProSerArg----	123
Db	17957	GCTTCGACGAGTAATTGGCCACCGCAGTCTCCGCTCGCACACATCGAAAGGGCGG	17898	
Oy		124	-----ProGluProAnaHrcYsgLysSerCYsaGlyIsaArgValGlyLeuthrClph	141
Db	17897	TCAATCCCCCTCCACTGGTGGTTCAGGTCGATGCTCATGCTTCATGTTCGGAAC	17838	
Oy		141	eLysCyS-----ArgCyGlinAsnleu	148
Db	17837	ACGTAGCTGACCGTATCGTGGCGGAAGTCTC	17807	

```

RESULT 11
US-11-186-731-4/c
; Sequence 4, Application US/11186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MP12001-047P1RCP1(M)
; CURRENT APPLICATION NUMBER: US/11/186,731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 24120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(71)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)...(23978)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (23979)...(24120)
US-11-186-731-4

Alignment Scores:
Pred. No.:          3.59e+03
Score:              86.00
Percent Similarity: 37.40%
Best Local Similarity: 28.24%
Query Match:        8.74%
GB:
              7
Gaps:         3
Length:       24120
Matches:      37
Conservative: 12
Mismatch:     71
Indels:       11

```

[illegible]

```

Db      CTGATGACAGCGCTTGAGGAGTGTGCCCGGACAGCGGACAGGACCTTGAATG
18208

QY      69  ALAAlaValAlGlnPro-ProAlaProValHisGluThrIlysLeuThrCys
      :::::
Db      18148  TCGGCTGTGCAGTGCAGCCAGCTCTCTCTTAAAGTCGGGGGGCCGACAA
      :::::
QY      88  gThrmelValProHisGlnSerSerSerTyrglnGlnAsp-----
      |||:::||||
Db      18088  AGAGCGCTGTGATGCGACAGATTCCTCTTCAACCCAGAGCTTGATTA
      |||:::||||
QY      105  rProAlaAlaAlaAlaProGlnAlaValIlysserSerIleAlaIlePro
      |||:::||||
Db      18028  GCGTGCACACAGGTACTTGGCGACCGAGTCTCCGCTCTGTGCCACACT
      |||:::||||
QY      124  -----ProGlnProAsnArgCysAsgIlySerCysArgIlyAsrValGly
      |||:::||||
Db      17968  TCATCCCCCTCCACCTGCTGCTTCAGTGCATCTGTCTGACGTTCAATCA
      |||:::||||
QY      141  eIlysCys-----ArgCysGlyAsnIleu 148
      :::::
Db      17908  ACGTAGTGAACGGTATTCGGTGGGGGAGTCTC 17878
      :::::

```

```

RESULT 12
US-11-087-100-21
; Sequence 21, Application US/11087100
; Publication No. US2005026640A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Platt, James
; APPLICANT: Kunert, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORF of a
; TITLE OF INVENTION: System and Uses Thereof
; FILE REFERENCE: 2997-23
; CURRENT APPLICATION NUMBER: US/11/087,100
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: 09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/284,066
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/298,796
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/323,269
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Schizochytrium sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
US-11-087-100-21

```

Alignment Scores:	123	length:	1323
Pred. No.:	83.00	Matches:	50
Score:	35.29%	Conservative:	27
Percent Similarity:	35.29%	Mismatches:	69
Best Local Similarity:	25.51%	Indels:	50
Query Match:	8.64%	Gaps:	10
DB:	7		
US-10-716-089-20 (1-188)	x	US-11-087-100-21 (1-1323)	

```
QY      10 ThrThrSerGlnAlaProIleuGlyPro-----ValMetCysLysA  
       ::::|||||  
Db      23 TGACTTCAGCGGCCTCCGGACGCCCATATGACCCTTAAGAAGACATCTCCT  
       |||||  
QY      27 PhePheGlySerGlnAlaThrMetGlyLeuCysSerIscyTyraArgG  
       :|||:  
Db      293 TTCTGGCCGTCAACCATATGACCGGCGCATCTCTCG-----  
       |||||
```

```

GENERAL INFORMATION:
APPLICANT: Metz, James
APPLICANT: Barclay, William
APPLICANT: Flact, James
APPLICANT: Kueer, Jerry
TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORF1 of a PUFA Polyketide Synthase
TITLE OF INVENTION: System and Uses Thereof
FILE REFERENCE: 2997-29
CURRENT APPLICATION NUMBER: US/11/087,100
CURRENT FILING DATE: 2005-03-21
PRIORITY APPLICATION NUMBER: 09/231,899
PRIORITY FILING DATE: 1999-01-14
PRIORITY APPLICATION NUMBER: 60/284,066
PRIORITY FILING DATE: 2001-04-16
PRIORITY APPLICATION NUMBER: 60/298,796
PRIORITY FILING DATE: 2001-06-15
PRIORITY APPLICATION NUMBER: 60/323,269
PRIORITY FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.3
SEQ ID NO 3
LENGTH: 6177
TYPE: DNA
ORGANISM: Schizochytrium sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(6177)
IS-11-087-100-3

```

US-10-716-089-20 (1-188) x US-11-087-100-3 (1-6177)		
QY	10 ThrThSerGlnAlaProGluGlyPro-----ValMetCysIysV	26
Db	1610 TCGAGCTTCAGCGCCTCCGACGCGCCATGACCCCTGAAAGACATGCTCT	1669
QY	27 PhePheGlySerGlnAlaThrMetGlyLeuCysSerIysCysTyrArg	46
Db	1670 TTCTGCGCTCACCAACCATTCAGCGCGCCATCTCTG-----	1711
QY	47 GlnAlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaGln	66
Db	1712 GAATGAAAAAGGAGGGAATGTCGCGCTTGTGCGGCTCGGACAG	1771
QY	67 Thr-----AlaAlaAlaValGln	76
Db	1772 ACCGTACCGCTGCTCGCGCTCTCAAGAGGCGCGCTCCCTGAAGC	1825
QY	77 ValHisGluThrIlybLeuThr---CysGluValGluArgThrMetIleV	95
Db	1826 -----AGAACTCAATGATCATGATCAGATCACTTAACGACTCGGCAC	1879
QY	96 SerSerSerTyrGlnGlnAspLeuValThr-----	105
Db	1880 ACA-----CCTCGTACATTGGCAACTCTCGGACGCGGTCTC	1930
QY	106 -----ProAlaAlaAlaAlaProGlnAlaValIysSerI	121
Db	1931 GCTTCACGGCGCCCTCTTACGATCACTCAGGCGAAACATCCGCTCA	1990
QY	122 Ser-Arg-----ProGluProAsnArgCysGlySerCysArgIlybArg	139
Db	1991 AGCTTCGCAAGTACCTCTCTGAGACCGCGAGGTTCATGCGCTGCTG	2050
QY	139 rGlyPheIysCysArgCys-GlyAsnLeuTyrCysAlaLeuHisArgTyr	155
Db	2051 ATCTGTCGG---CAGTGGCGAAAACTTTAAGTCAAGTCTGCGCGCTT	2107
QY	156 --SerAspIysHisThrCysThrTyrAspTyrIlybAlaAlaGly	169
Db	2108 CTTCCGATACCCCGCGCGCCAGCTTTGACCGCCGCGCGATGGC	2151
RESULT 14		
US-10-821-234-165/C		
Sequence 165, Application US/10821234		
Publication No. US20050255114A1		
GENERAL INFORMATION:		
APPLICANT: Labat, Ivan		
APPLICANT: Seache-Crain, Birgit		
APPLICANT: Andarmari, Susan		
APPLICANT: Tang, Y. Tom		
TITLE OF INVENTION: Methods for diagnosis and Treatment of P		
FILE REFERENCE: 821A		
CURRENT APPLICATION NUMBER: US/10/821,234		
CURRENT FILING DATE: 2004-04-07		
PRIOR APPLICATION NUMBER: US 60/462,047		
PRIOR FILING DATE: 2003-04-07		
NUMBER OF SEQ ID NOS: 1704		
SOFTWARE: pc_seq_genes Version 1.0		
SEQ ID NO 165		
LENGTH: 2815		
TYPE: DNA		
ORGANISM: Homo sapiens		
US-10-821-234-165		
Alignment Scores:		
Pred. No.:	489	Length: 2815
Score:	83.00	Matches: 40
Percent Similarity:	37.06%	Conservative: 23
Best Local Similarity:	25.29%	Mismatches: 66
Query Match:	8.43%	Indels: 41
GB:	6	Gaps: 8

US-10-716-089-20 (1-188) x US-10-821-234-165 (1-2815)

```

QY 22 LysaAnLeuCyGlyPhePheGlySerGlnAlaThrMetMetGlyLeuCySerIleCySerTyr 41
   |||||
Db 1802 AAAAACATTGCTGGTTCAGAGTCCACAGACACACCTCAAGGCTCTGGGGAGGCC--- 1746
   |||||
QY 42 ArgGluThrValMetGlnAlaLysMetThrAla-----LeuAlaGluGlnAla 57
   |||||
Db 1745 -----CGATGCTCAATGCCAGCATGTGGCATCTTCGTGAGACTACCGCTGGCCAGC 1692
   |||||
QY 58 ThrGlnAlaAlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProVal 77
   |||||
Db 1691 ATTGTCCTCACTTCATGCTCCAGCTCCAGCTGGGTAGGCTGCAGACCTCAACCCGACATC 1632
   |||||
QY 78 HisGluThr-----LysLeuThrCySerGluValGluArgThrMetIleVal-ProHisGlu 95
   |||||
Db 1631 CAGACCACTCCCTCCACAGACACTTGTAGCCCGAGCTGCTCATGCTCTTGATTCAGACA 1572
   |||||
QY 95 nSerSerSerTyrGlnGlnAlaPheLeuValThrProAlaAlaAlaProGlnAlaValIly 115
   |||||
Db 1571 AAGT-----GTGCCAGACAGCCCATGACCTCGTACTAAG 1536
   |||||
QY 115 sSerSerIle-----AlaAlaProSerArgProGluProAlaArgCySerIle 131
   |||||
Db 1535 AAAGAGCTTGTGGCGCATTTGAAGTGACATCAGCGGCTGACCTTGTGGAGCCAGC 1476
   |||||
QY 131 rCyArgArgValArgValGlyLeuThrGlyPheLys----- 142
   |||||
Db 1475 TGTGAGGCGAGAGACAGGGGTGACAAATTGTCTTGGCAGTGTACTCCAGCTGTGGC 1416
   |||||
QY 143 -CyArgCySv-----GlyAlaLeuTyrCyValAlaLeu 153
   |||||
Db 1415 ATGCAGGTGTGACGAAGTATGATCAGCAATGAGGGGTGTGCAATATCTGTGAGCTCA 1356
   |||||
QY 153 sArgTyrSerArgLysHisThrCySerThr 162
   |||||
Db 1355 TCACCAACAGCGGAAAAAGG---TGCACCT 1331

```

RESULT 15
US-10-508-263-103

```

/ Sequence 103, Application US/10508263
/ Publication No. US20050260754A1
/ GENERAL INFORMATION:
/ APPLICANT: BASF Plant Science GmbH
/ TITLE OF INVENTION: Constructs and methods for regulating gene expression
/ FILE REFERENCE: 53262-20085.00
/ CURRENT APPLICATION NUMBER: US/10/508,263
/ NUMBER OF SEQ ID NOS: 126
/ SOFTWARE: PatencIn Ver. 2.1
/ SEQ ID NO 103
/ LENGTH: 801
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(798)
/ OTHER INFORMATION: 22kd alpha-zelin
US-10-508-263-103

```

Alignment Scores:

```

Pred. No.: 115 Length: 801
Score: 82.50 Matches: 48
Percent Similarity: 29.89% Conservative: 7
Best Local Similarity: 26.09% Mismatches: 91
Query Match: 8.38% Indels: 38
DB: 6 Gaps: 6

```

US-10-716-089-20 (1-188) x US-10-508-263-103 (1-801)

```

QY 32 AlaThrMetGlyLeuCySerIleCySerTyrArgGluThrValMetGlnAlaLysMetThr 51
   |||||
Db 168 GCTACAAACAGCGCTTGGCGAGCGCTTACACAAACCGTTTGGCCATTTACAAACACA 227
   |||||

```

```

QY 52 AlaLeu---AlaGluGlnAlaThrGlnAlaAlaGlnAlaAlaThrSerAlaT 70
   |||||
Db 228 ATCTTGGCACATCTTAACCATTAACCAATCGCAACACAACTAGGCAAC 287
   |||||
QY 71 ValGlnProProAlaProValHisGluThrLysLeuThrCySerGluValG 90
   |||||
Db 288 CGCATTTAGCCCAATAGCGCGGTGAACTGTGTCTTACTTGCACAAAGC 347
   |||||
QY 91 IleValProHisGlnSerSerSerTyrGlnGlnAlaPheLeuValThrPro 110
   |||||
Db 348 ATCCAAACCACTTGTCTTGGCGAACAACAGCCGC-----ATACAG 398
   |||||
QY 111 ProGlnAlaValLysSerSerIleAlaAlaProSerArgProGlu----- 125
   |||||
Db 399 GTTGCAACAGTTTCTACAGCTCTTATGCAACTAGCCAGGGGAAACCT 458
   |||||
QY 126 -----ProAlaArgCySerIle 132
   |||||
Db 459 GCAACAGCAACACTGTTTCATCTAGCCCACTGCTGTGGCAATGCG 518
   |||||
QY 133 -----ArgLysArgValGlyLeuThrGlyPheLys 146
   |||||
Db 519 GCAACAGCACTGTCAACAGATCTGACCGGC-----TCTTAGTCAG 572
   |||||
QY 147 AsnLeuTyrCyValAlaLeuHisArgTyrSerAsp----- 159
   |||||
Db 573 GAACCTTACTGCTTACTTAACAAGCTTCTTCATTCACCAACTAGAT 632
   |||||
QY 160 ThrCyThrTyrAspTyrLysAlaAlaGlyGln-GluAlaIleAlaLys 179
   |||||
Db 633 TGCTGCTGACTTACAAACAGCGGACAGACTATTAACTCACTTGACAGCG 692
   |||||
QY 179 uValValAla 182
   |||||
Db 693 GGTTCGCGCGC 702

```

Search completed: December 8, 2005, 19:02:26
Job time : 265 secs